

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:12:44 ; Search time 200 Seconds  
(without alignments)  
2926.263 Million cell updates/sec

Title: US-10-736-769-4  
Perfect score: 6909  
Sequence: 1 MAAGAGRGWLLWALLRLAQ.....GSIKAGAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	8	ADJ27297
2	6909	100.0	1332	9	ADY60740 Human NPC
3	6909	100.0	1332	9	AEB93570 Human NPC
4	6896	99.8	1332	4	AAM79169 Human pro
5	6896	99.8	1332	4	AAG5638
6	6872.5	99.5	1359	4	AAM79168
7	6872.5	99.5	1359	4	ADJ27337
8	6872.5	99.5	1359	9	ADY60780 Human NPC
9	6872.5	99.5	1359	9	AEB93610 Human NPC
10	6536	94.6	1344	4	ABG22693
11	5421.5	78.5	1331	8	ADJ27295
12	5421.5	78.5	1331	9	ADY60738
13	5421.5	78.5	1331	9	AEB93568
14	5407	78.3	1333	8	ADJ27305
15	5407	78.3	1333	9	ADY60748
16	5407	78.3	1333	9	AEB93578
17	4466	64.6	982	4	ABG22691
18	2402.5	34.8	1278	3	AAW88445
19	2402.5	34.8	1278	3	AAB42983
20	2402.5	34.8	1278	8	ADQ39879
21	2402.5	34.8	1278	8	ADU06723
22	2385	34.5	1319	2	AAW88446
23	2262	32.7	1287	4	ABB61737
24	2262	32.7	1287	8	ADS96670 Drosophil

25	1864.5	27.0	1223	4	ABB58629
26	1329	19.2	1170	2	AAW88447
27	1078	15.6	1296	2	AAW88448
28	963	13.9	229	4	AAM80153
29	963	13.9	229	4	AAM80152
30	942	13.6	194	4	AAG5637
31	651.5	9.4	1447	2	AAW75375
32	651.5	9.4	1447	2	AAW52200
33	651.5	9.4	1447	2	AAW72969
34	651.5	9.4	1447	4	AAB67163
35	651.5	9.4	1447	5	AAE19830
36	651.5	9.4	1447	5	ABJ10931
37	651.5	9.4	1447	5	AAG79571
38	651.5	9.4	1447	7	ABU62275
39	651.5	9.4	1447	7	ADD46678
40	651.5	9.4	1447	7	ADE94224
41	651.5	9.4	1447	7	ADH62731
42	651.5	9.4	1447	8	ADH48989
43	646.5	9.4	1434	2	AAW52199
44	646.5	9.4	1434	2	AAW72968
45	646.5	9.4	1434	4	AAB67159

ALIGNMENTS

RESULT 1

ADJ27297

ID ADJ27297 standard; protein; 1332 AA.

XX AC ADJ27297;

XX AC ADJ27297;

DT 20-MAY-2004 (first entry)

XX DE Human NPC1L1.

XX DE Human NPC1L1.

XX KW Niemann-Pick disease; type Cl; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promoter;  
KW sterol regulated element binding protein 1; SREBP1;  
KW binding consensus sequence; transmembrane domain; sterol-sensing domain; SSD; cholesterol; NPC1; receptor; Niemann-Pick Cl disease; intestinal;  
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;  
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX OS Homo sapiens.

XX PN WO2004009772-A2.

XX PD 29-JAN-2004.

XX PF 17-JUL-2003; 2003WO-US022467.

XX PR 19-JUL-2002; 2002US-0397442P.

XX PA (SCHE ) SCHERING CORP.

XX PI Altman SW, Murgolo NJ, Wang LQ, Graziano MP;

XX DR WPI; 2004-132945/13.

XX DR N-PSDB; ADJ27296.

XX DR GENBANK; AF192522.

XX PT New Niemann-Pick disease, type Cl, gene-like 1 polypeptide, for detecting or identifying antagonists of NPC1L1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

XX PS Claim 11; SEQ ID NO 4; 125pp; English.

XX CC This sequence represents a Niemann-Pick disease, type Cl, gene-like 1 (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and

CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a  
CC sterol regulated element binding protein 1 (SREBP1) binding consensus  
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-  
CC sensing domain (SSD) which is involved in sensing cholesterol levels,  
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1  
CC has 42% amino acid sequence homology to human NPC1, a receptor  
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or  
CC polynucleotides are useful for detecting or identifying antagonists of  
CC NPC1L1, which can be used for inhibiting intestinal cholesterol  
CC absorption in a subject, or for treating medical conditions including  
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary  
CC heart disease, stroke or arteriosclerosis.

xx  
SQ Sequence 1332 AA;

Query Match	100.0%;	Score 6909;	DB 8;	Length 1332;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1332;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MAEAGLRGWLWALLRLAQAQSEPTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60		
Db 1	MAEAGLRGWLWALLRLAQAQSEPTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60		
QY 61	TPARKITGDHLILLOKICPRLYTGENTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db 61	TPARKITGDHLILLOKICPRLYTGENTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
QY 121	VNLHCHNTCSNPQSLFVNTVAQLGAGQLPAVVAEYAFQHSFAEQSYDCSRRVPA	180		
Db 121	VNLHCHNTCSNPQSLFVNTVAQLGAGQLPAVVAEYAFQHSFAEQSYDCSRRVPA	180		
QY 181	ATLAVGTGCVGVGSMALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIOPANEV	240		
Db 181	ATLAVGTGCVGVGSMALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIOPANEV	240		
QY 241	ARCNSQGDVATCSQDCQCAAPARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300		
Db 241	ARCNSQGDVATCSQDCQCAAPARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300		
QY 301	LLVGRFVPAPDKSKMDPKGTSLSDKLSFSTHTLLGQFFQGWCTWVASMPLTILVLSV	360		
Db 301	LLVGRFVPAPDKSKMDPKGTSLSDKLSFSTHTLLGQFFQGWCTWVASMPLTILVLSV	360		
QY 361	IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHEGPFERTNQVILTAPNRSY	420		
Db 361	IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHEGPFERTNQVILTAPNRSY	420		
QY 421	RYDSLGLGPKNFGSLDLDLLELLELQERLRLQVWSPQAQRNLSLQDICVAPLNPDNT	480		
Db 421	RYDSLGLGPKNFGSLDLDLLELLELQERLRLQVWSPQAQRNLSLQDICVAPLNPDNT	480		
QY 481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCNAAPLTFKDGTALAL	540		
Db 481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCNAAPLTFKDGTALAL	540		
QY 541	SCWADYGAPVFPPLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM	600		
Db 541	SCWADYGAPVFPPLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM	600		
QY 601	RAFQRRMAGMFQVFTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV	660		
Db 601	RAFQRRMAGMFQVFTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV	660		
QY 661	MVDSKATLGLGVAVVLGVAAMAGPFSYIGIRSSLVILQVVPFLVSVGADNIRIFVLE	720		
Db 661	MVDSKATLGLGVAVVLGVAAMAGPFSYIGIRSSLVILQVVPFLVSVGADNIRIFVLE	720		
QY 721	YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTFALTSGLA	780		
Db 721	YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTFALTSGLA	780		
QY 781	ILDPELLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840		

Db	781	ILDPELLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840
QY	841	LHWITRGVLLFLALFGVLSYSMCHIISVGLDOELALPKDSYLLDYFELFNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVLSYSMCHIISVGLDOELALPKDSYLLDYFELFNRYFEVGAPV	900
QY	901	YFVTTLGYNFSSEAGNNAICSSAGCNFNFSFTQKIQVATEFPFQSYLAIPASSWVDDFIDW	960
Db	901	YFVTTLGYNFSSEAGNNAICSSAGCNFNFSFTQKIQVATEFPFQSYLAIPASSWVDDFIDW	960
QY	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCIKNCMSITMGSVRPSVEQFHKYLFWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCIKNCMSITMGSVRPSVEQFHKYLFWFLNDRP	1020
QY	1021	NITKCPKGLAAAYSTSVNLTSDGOVLASRFMAVYHKLKNSODVTEALRAARELAANITADL	1080
Db	1021	NITKCPKGLAAAYSTSVNLTSDGOVLASRFMAVYHKLKNSODVTEALRAARELAANITADL	1080
QY	1081	RKVPGTDPAPFEVPPYITITNVFYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140
Db	1081	RKVPGTDPAPFEVPPYITITNVFYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140
QY	1141	NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVCMSEVFVSHITRSFAISTKPTWLER	1200
Db	1141	NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVCMSEVFVSHITRSFAISTKPTWLER	1200
QY	1201	AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIIFFRMLLITLGLLHGLVFLPV	1260
Db	1201	AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIIFFRMLLITLGLLHGLVFLPV	1260
QY	1261	ILSYVGPDPVNPALALBQKRAEAAVAAVMVASCNPHSPRSVSTADNIYVNHSEFSGIKGAGA	1320
Db	1261	ILSYVGPDPVNPALALBQKRAEAAVAAVMVASCNPHSPRSVSTADNIYVNHSEFSGIKGAGA	1320
QY	1321	ISNPLPNNGRQF 1332	
Db	1321	ISNPLPNNGRQF 1332	
RESULT 2			
ADY60740	ID	ADY60740 standard; protein; 1332 AA.	
XX	AC	ADY60740;	
XX	DT	19-MAY-2005 (first entry)	
XX	DE	Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 4.	
XX	KW	Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;	
XX	KW	Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;	
XX	KW	metabolic disorder; atherosclerosis; cardiovascular disease;	
XX	KW	coronary artery disease; cerebrovascular ischemia; neurological disease;	
XX	KW	arteriosclerosis; transgenic animal.	
XX	OS	Homo sapiens.	
XX	PN	WO2005015988-A1.	
XX	PD	24-FEB-2005.	
XX	PF	16-DEC-2003; 2003WO-US040113.	
XX	PR	17-JUL-2003; 2003US-00621758.	
XX	PR	22-AUG-2003; 2003US-00646301.	
XX	PR	16-SEP-2003; 2003US-00663208.	
XX	PA	(SCHE ) SCHERING CORP.	
XX	PI	Altman SW, Murgolo NJ, Wang LQ, Graziano MP;	
XX	DR	WPI; 2005-284403/29.	
XX	DR	N-PSDB; ADY60739, ADY60787.	

XX Novel isolated NPC1l1 polypeptide, useful for identifying antagonist of  
PT NPC1l1 that inhibits NPC1l1 mediated sterol or 5 alpha-stanol uptake, for  
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.  
XX Claim 30; SEQ ID NO 4; 146pp; English.  
XX  
CC The invention relates to an isolated Niemann-Pick C1-like protein 1  
CC (NPC1l1) polypeptide comprising 42 or more contiguous amino acids from an  
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,  
CC human and Mouse NPC1l1). Also included are an isolated polynucleotide  
CC encoding NPC1l1 (appearing as ADY60737, ADY60739 or ADY60747), a  
CC recombinant vector comprising the polynucleotide, a host cell comprising  
CC the vector, an isolated antibody which specifically binds to NPC1l1 (or  
CC to a rat NPC1l1 peptide sequence appearing as ADY60775-ADY60778),  
CC preparation of NPC1l1, a mutant transgenic mouse comprising a homozygous  
CC mutation of endogenous chromosomal NPC1l1 (where the mouse does not  
CC produce any functional NPC1l1 protein), an offspring or progeny of the  
CC mouse having inherited a mutated NPC1l1 allele, an offspring or progeny of the  
CC sample for an intestinal sterol or Salpha-stanol absorption, screening a  
CC inhibiting NPC1l1 mediated sterol or Salpha-stanol absorption antagonist,  
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical  
CC dosage form, and information indicating that NPC1l1 is a target of  
CC ezetimibe or substituted azetidinone), decreasing the level of intestinal  
CC sterol or Salpha-stanol absorption in a subject (involving reducing the  
CC level of expression of NPC1l1 in the subject), identifying an antagonist  
CC of NPC1l1 and an isolated mammalian cell which lacks a gene encoding a  
CC functional NPC1l1 protein. NPC1l1 is useful for identifying an antagonist  
CC of NPC1l1 which is useful for inhibiting or decreasing the level of  
CC NPC1l1 mediated sterol or Salpha-stanol uptake in a subject such as  
CC mouse, rat or human. The level of expression of NPC1l1 in the subject is  
CC reduced by mutating NPC1l1 in the subject. The NPC1l1 knockout mouse is  
CC useful for screening a sample for intestinal sterol or Salpha-stanol  
CC absorption antagonist. The NPC1l1 antagonists identified are useful for  
CC treating medical conditions caused or mediated by NPC1l1, e.g.,  
CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or  
CC arteriosclerosis. The present sequence represents a mammalian NPC1l1  
XX protein.  
XX  
SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 9; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAEAGLGRWLLWALLRLAQSEPTTHIQGYCAFYDECGKPELSGLMTLSNVSCLSN 60  
DB 1 MAEAGLGRWLLWALLRLAQSEPTTHIQGYCAFYDECGKPELSGLMTLSNVSCLSN 60  
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120  
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120  
QY 121 VNLHCHTCSNQSLFINTRVLAQLGAGQLPAVVAEAFYQHSFAPQSDSCSRVRVPA 180  
DB 121 VNLHCHTCSNQSLFINTRVLAQLGAGQLPAVVAEAFYQHSFAPQSDSCSRVRVPA 180  
QY 181 ATLAVGTMCGVYGSALCNQRMWLFQDGTGNGLAPLDITPHLLERQAVGSGIQPLNEGV 240  
DB 181 ATLAVGTMCGVYGSALCNQRMWLFQDGTGNGLAPLDITPHLLERQAVGSGIQPLNEGV 240  
QY 241 ARCNEQSGDDVATCSQDCAACPAJARPAQALDSTFYLQMPGSLVLIILCSFAVAVTI 300  
DB 241 ARCNEQSGDDVATCSQDCAACPAJARPAQALDSTFYLQMPGSLVLIILCSFAVAVTI 300  
QY 301 LLVGRFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFPQGWGTWASWPLTILVLSV 360  
DB 301 LLVGRFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFPQGWGTWASWPLTILVLSV 360  
QY 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFPHDQFGPFRTNQVILTAPNRSSY 420  
DB 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFPHDQFGPFRTNQVILTAPNRSSY 420

QY 421 RYDSSLGPKNFGSILDLDLLELLELQERLRLHQLVMSPEAQRNLSIQDICYAPLNPNT 480  
DB 421 RYDSSLGPKNFGSILDLDLLELLELQERLRLHQLVMSPEAQRNLSIQDICYAPLNPNT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFDGTALAL 540  
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFDGTALAL 540  
QY 541 SCHADYGAPVFPPLAIGGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLWEAEFLHEM 600  
DB 541 SCHADYGAPVFPPLAIGGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLWEAEFLHEM 600  
QY 601 RAPQRMAGHPQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGYSYSSRV 660  
DB 601 RAPQRMAGHPQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGYSYSSRV 660  
QY 661 MYDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSVGADNIFIFVLE 720  
DB 661 MYDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSVGADNIFIFVLE 720  
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFFLGALTMPAVRTFALTSGSLAV 780  
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFFLGALTMPAVRTFALTSGSLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDYCCVKPQELPPPGQEGELLGFFQKAYAPFL 840  
DB 781 ILDFLLQMSAFVALLSLDSKQEAASRLDYCCVKPQELPPPGQEGELLGFFQKAYAPFL 840  
QY 841 LHWITRGVLLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
DB 841 LHWITRGVLLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSFTQIOVATEFPQSQYLAIIPASSWVDDFIDW 960  
DB 901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSFTQIOVATEFPQSQYLAIIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRPSVQFHKYLPWFNDLRP 1020  
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRPSVQFHKYLPWFNDLRP 1020  
QY 1021 NIKCPKGLAAYSTSNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELAANTADL 1080  
DB 1021 NIKCPKGLAAYSTSNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELAANTADL 1080  
QY 1081 RKVPGTDPAPFVPPYITTNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGL 1140  
DB 1081 RKVPGTDPAPFVPPYITTNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGL 1140  
QY 1141 NLLSVIMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAI STKPTWLER 1200  
DB 1141 NLLSVIMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAI STKPTWLER 1200  
QY 1201 AKEATISMGSAVPAVGMNTNLPGLIVLGLAKAQLQIPIFFRNLNLIITLGLLHGLVFLPV 1260  
DB 1201 AKEATISMGSAVPAVGMNTNLPGLIVLGLAKAQLQIPIFFRNLNLIITLGLLHGLVFLPV 1260  
QY 1261 ILSYVGPDPNVPALAEQKRAEAAVAVMVASCPNHPSPRVSTADNLYVNHSPSGSIKGAGA 1320  
DB 1261 ILSYVGPDPNVPALAEQKRAEAAVAVMVASCPNHPSPRVSTADNLYVNHSPSGSIKGAGA 1320  
QY 1321 ISNFLPNNGRQF 1332  
DB 1321 ISNFLPNNGRQF 1332  
RESULT 3  
AEB93570  
ID AEB93570 standard; protein; 1332 AA.  
XX  
AC AEB93570;  
XX  
DT 06-OCT-2005 (first entry)  
XX

Human NPC1L1 protein.

protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic; cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke; arteriosclerosis; human.

**Homo sapiens.**

WC2005069900-A2.

04-AUG-2005.

14--JAN-2005: 2005WO-US001469.

16-TAN-2004: 2004US-0537341P.

(MERT) MERTCK &amp; CO. INC.

Demetrius C. J. M.

REF 2005 EC4070/57

N-PSDB; AEB93569.

## Identifying Irgain

glucuronide an

— 1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-33-34-35-36-37-38-39-40-41-42-43-44-45-46-47-48-49-50-51-52-53-54-55-56-57-58-59-60-61-62-63-64-65-66-67-68-69-70-71-72-73-74-75-76-77-78-79-80-81-82-83-84-85-86-87-88-89-90-91-92-93-94-95-96-97-98-99-100-101-102-103-104-105-106-107-108-109-110-111-112-113-114-115-116-117-118-119-120-121-122-123-124-125-126-127-128-129-130-131-132-133-134-135-136-137-138-139-140-141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-160-161-162-163-164-165-166-167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000-1001-1002-1003-1004-1005-1006-1007-1008-1009-1010-1011-1012-1013-1014-1015-1016-1017-1018-1019-1020-1021-1022-1023-1024-1025-1026-1027-1028-1029-1030-1031-1032-1033-1034-1035-1036-1037-1038-1039-1040

The invention relates to identifying a ligand of NPC1L1. The method involves contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC1L1. In identifying a ligand of NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands are useful for stimulating or blocking the activity of NPC1L1, and for treating conditions caused or mediated by NPC1L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents a human NPC1L1, a N-glycosylated protein.

Sequence 1332 AA:

Query Match	Score 6909;	DB 9;	Length 1332;
Best Local Similarity	100.0%;	Pred. No. 0;	
Best Match 1332; Conservative	100.0%;	0; Mismatches	0; Indels
		0; Mismatches	0; Gaps

1 MAEAGT.RGWT.LWALLRLAQSEPYTTHQPGYCAFYDECCKNPELSGSLMTLSNVSLTN 60

1 MAFAGI.RGW.I.WALLRLAOSEPYTTHQPGYCAFYDECGKNPELGSIMTLNVCLSN 60

61 TRAPKITGDHI.IILOKICPRLYTGPNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120

C1 M000VITC0DHI.II.I.OKTCPRLYTGPNT0ACCSAKQLVSLEASLSITKALLTRCPACSDNF 120

101 187 HCNHHCSPNOSI.FTNVTRVAOLGAGOLPAWAYEAFYQHSAEQSYDSCSRVRPAA 180

180 VPSRVRSCSYDQSFHAYEAVAGOLPVAEYQHSFAECSRVVPAA 180

[illegible][illegible]

... TCTTCTCCCTTATCTCCCTCCCAACCCPATARPQALDSTFYLGOMPGSLVLIILCSVFAVVTI 300

... - - - - - TCSCCNCDCRASCPTARPOALDSTFFYLGOMPGSLVLIILCSVFAVVTI 300

[illegible]

-----SDKT.SFSTHTLLGOFFOQGWTVASWPLTILVLSV 360

361	QY	IPVVALAAGLVFTELTTPDVELASAPNSQARSEKAFHQDFGFPFKINUVILAFARSGT	420
361	DB	IPVVALAAGLVFTELTTPDVELASAPNSQARSEKAFHQDFGFPFKINUVILAFARSGT	420
421	QY	RYDSLGLGPKNFSGILDLDDLLELLEQLRHLQVWSPEAQRNISLQDICIYAPLNPDNT	480
421	DB	RYDSLGLGPKNFSGILDLDDLLELLEQLRHLQVWSPEAQRNISLQDICIYAPLNPDNT	480
481	QY	SLYDCCINSLLQYFONNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGDTALAL	540
481	DB	SLYDCCINSLLQYFONNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGDTALAL	540
541	QY	SCMADYGAVFPFLAIGGYKGKQYSEABALIMTFSLNYPAGDPRLAQAKLWEAFLEEM	600
541	DB	SCMADYGAVFPFLAIGGYKGKQYSEABALIMTFSLNYPAGDPRLAQAKLWEAFLEEM	600
601	QY	RAFQRMAGMFQVTFTAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGSSYSSSRV	660
601	DB	RAFQRMAGMFQVTFTAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGSSYSSSRV	660
661	QY	MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLIVLSVGADNIFIFVLE	720
661	DB	MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLIVLSVGADNIFIFVLE	720
721	QY	YQRLPRRGPBPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTPMPAVRFTALTSGIYAV	780
721	DB	YQRLPRRGPBPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTPMPAVRFTALTSGIYAV	780
781	QY	ILDPELLQMSAFVALLSLDSKQBSASRLDVCCKVKPQELPPQGGSGLLGLGFFQKAYAPFL	840
781	DB	ILDPELLQMSAFVALLSLDSKQBSASRLDVCCKVKPQELPPQGGSGLLGLGFFQKAYAPFL	840
841	QY	LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGAPV	900
841	DB	LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGAPV	900
901	QY	YFVTTLGHNSSEAGMNAICSSAGCNFSPQIKIQTATEFFPQGSYLAIPASSWVDDFIDW	960
901	DB	YFVTTLGHNSSEAGMNAICSSAGCNFSPQIKIQTATEFFPQGSYLAIPASSWVDDFIDW	960
961	QY	LTPSSCCRLYISGPNKDKFCPSTVNSLNCILKNCMSITMGSVRPSVQFHXYLPWFLNDRP	1020
961	DB	LTPSSCCRLYISGPNKDKFCPSTVNSLNCILKNCMSITMGSVRPSVQFHXYLPWFLNDRP	1020
1021	QY	NIKCPKGGIAYSTSVNLTSQGVLASRFWAYHKPLKNSQDYTEALARAEALAAANTADL	1080
1021	DB	NIKCPKGGIAYSTSVNLTSQGVLASRFWAYHKPLKNSQDYTEALARAEALAAANTADL	1080
1081	QY	RKVPGTDPAFEFVPTYITINVEYQVLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL	1140
1081	DB	RKVPGTDPAFEFVPTYITINVEYQVLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL	1140
1141	QY	NLLSIVMILVDVTGFMALWDISYNAVSLINLVSAGVSVFVSHITRSFAISTKPTWLER	1200
1141	DB	NLLSIVMILVDVTGFMALWDISYNAVSLINLVSAGVSVFVSHITRSFAISTKPTWLER	1200
1201	QY	AKBATISMGSAFVAGVAMTNLPGIILVLGAKAQLIQIFFFFRLNLIITLGLLHGLVPLPV	1260
1201	DB	AKBATISMGSAFVAGVAMTNLPGIILVLGAKAQLIQIFFFFRLNLIITLGLLHGLVPLPV	1260
1261	QY	ILSVGPDVNPALALQKAEAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPGSGIKGAGA	1320
1261	DB	ILSVGPDVNPALALQKAEAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPGSGIKGAGA	1320
1321	QY	ISNPLPNNGRQF 1332	
1321	DB	ISNPLPNNGRQF 1332	
RESULT 4			
AAM79169			
ID AAM79169 standard; protein; 1332 AA.			
XX			

RESULT 4  
AAM79169  
ID AAM79169 standard; protein; 1332 AA.  
XX

AC	AAW79169;	
XX	06-NOV-2001 (first entry)	
DT	Human protein SEQ ID NO 1831.	
DE		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157190-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	05-FEB-2001; 2001WO-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Chen ZW;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
DR	WPI: 2001-476283/51.	
DR	N-PSDB: AAK52302.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
XX		
PS	Claim 20; Page 4216-4219; 6221pp; English.	
XX		
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the	
CC	encoded polypeptides (AAW78323-AAW80302) that exhibit activity relating to	
CC	cytokine, cell proliferation or cell differentiation or which may induce	
CC	production of other cytokines in other cell populations. The	
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	
CC	peptide therapy. The polypeptides have various cytokine-like activities,	
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	
CC	activity, tissue growth factor activity, immunomodulatory activity and	
CC	activin/inhibin activity and may be useful in the diagnosis and/or	
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111	
CC	(AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the	
CC	sequence listing were missing at the time of publication	
XX		
SQ	Sequence 1332 AA;	
	Query Match 99.8%; Score 6896; DB 4; Length 1332;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 MAEAGLRGWLWALLRLAQSPYTTIHOPGCAFYDECGKPNELSGSLMTLSNVCSLSN 60	
DB		
DB	1 MAEAGLRGWLWALLRLAQSPYTTIHOPGCAFYDECGKPNELSGSLMTLSNVCSLSN 60	
QY	61 TPARKITGDHLLILLOKICPRLVTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120	
DB		
DB	61 TPARKITGDHLLILLOKICPRLVTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120	
QY	121 VNLHCHNTCSNQSIFINVTVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180	
DB		
DB	121 VNLHCHNTCSNQSIFINVTVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180	

QY	181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGIAPLDITFHLLEPQAVGSGIQPLNEGV 240	
DB		
DB	181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGIAPLDITFHLLEPQAVGSGIQPLNEGV 240	
QY	241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPSGLVLIILCSVFAVVTI 300	
DB		
DB	241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPSGLVLIILCSVFAVVTI 300	
QY	301 LLVGRVAPARDKSKWVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360	
DB		
DB	301 LLVGRVAPARDKSKWVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360	
QY	361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHFQFFRTNQVILTAPNRSSV 420	
DB		
DB	361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHFQFFRTNQVILTAPNRSSV 420	
QY	421 RYDSLILGPKNFGILDLDLLELELOERLHLQWSPQAQRNLSLODICYAPLNPDT 480	
DB		
DB	421 RYDSLILGPKNFGILDLDLLELELOERLHLQWSPQAQRNLSLODICYAPLNPDT 480	
QY	481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDWKDHFLYCANAPLTFKDGTFALAL 540	
DB		
DB	481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDWKDHFLYCANAPLTFKDGTFALAL 540	
QY	541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600	
DB		
DB	541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600	
QY	601 RAFQRMAGMFOVTFEASRLDEINRTTAEDLPIFATSYIVIFLYSIALGSYSWSRV 660	
DB		
DB	601 RAFQRMAGMFOVTFEASRLDEINRTTAEDLPIFATSYIVIFLYSIALGSYSWSRV 660	
QY	661 MYDSKATILGLGGAVVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFLVLE 720	
DB		
DB	661 MYDSKATILGLGGAVVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFLVLE 720	
QY	721 YORLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGTALTPMPAVRTFALTSLAV 780	
DB		
DB	721 YORLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGTALTPMPAVRTFALTSLAV 780	
QY	781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQKQELPPQCGSGLLLGFFQKAYAPFL 840	
DB		
DB	781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQKQELPPQCGSGLLLGFFQKAYAPFL 840	
QY	841 LHWITRGVVLFLFLALFGVLSYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFEVGAPV 900	
DB		
DB	841 LHWITRGVVLFLFLALFGVLSYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFEVGAPV 900	
QY	901 YFVTTLGYNFSSEAGMNAICSSAGCNCNFSFTQKIQYATEFFPQSYLAIPASSWVDDFIDW 960	
DB		
DB	901 YFVTTLGYNFSSEAGMNAICSSAGCNCNFSFTQKIQYATEFFPQSYLAIPASSWVDDFIDW 960	
QY	961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNCSITMGSVRPSVEQHKYLPWFINDRP 1020	
DB		
DB	961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNCSITMGSVRPSVEQHKYLPWFINDRP 1020	
QY	1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAVHKPKNSQDYTEALRAARELAANTADL 1080	
DB		
DB	1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAVHKPKNSQDYTEALRAARELAANTADL 1080	
QY	1081 RKVPGTDPAFVFPYTTIINVFVEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140	
DB		
DB	1081 RKVPGTDPAFVFPYTTIINVFVEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140	
QY	1141 NLLSVIMLVDTVGPMAWDISYNVAVSLINLVSAVMSVFEVSHITRSFAISTKTWLER 1200	
DB		
DB	1141 NLLSVIMLVDTVGPMAWDISYNVAVSLINLVSAVMSVFEVSHITRSFAISTKTWLER 1200	
QY	1201 AKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFFFRNLNLLITLGLLHGVFLPV 1260	
DB		
DB	1201 AKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFFFRNLNLLITLGLLHGVFLPV 1260	



Db 1261 ILSYVGPVNPALALQKRAEEAAVMVASCPNHPRSVSTADNIYVNHSEFGSIKAGA 1320  
QY 1321 ISNFLPNNGRQF 1332  
Db 1321 ISNFLPNNGRQF 1332  
RESULT 6  
AAM79168  
ID AAM79168 standard; protein; 1359 AA.  
AC AAM79168;  
XX 06-NOV-2001 (first entry)  
XX Human protein SEQ ID NO 1830.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
XX WO200157190-A2.  
XX  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US004098.  
XX 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663561.  
PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.  
XX  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
DR N-P8DB; AAK52301.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
PT  
PT  
PS Claim 20; Page 4213-4216; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX  
XX Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVCLSN 60  
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVCLSN 60  
QY 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120  
QY 121 VNLHCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFQHSFABQSVDSCSRVPAA 180  
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFQHSFABQSVDSCSRVPAA 180  
QY 181 ATLAVGTMCGVYGVSALCNAQRWLNFGQDTGNGIAPLIDITFHLLPEQAVGSGIQPLNEGV 240  
Db 181 ATLAVGTMCGVYGVSALCNAQRWLNFGQDTGNGIAPLIDITFHLLPEQAVGSGIQPLNEGV 240  
QY 241 ARCNSQGDVATCSQDCAASCPIAIPQALDSTFYLGQMPGSLVLIILICSVFAVVTI 300  
Db 241 ARCNSQGDVATCSQDCAASCPIAIPQALDSTFYLGQMPGSLVLIILICSVFAVVTI 300  
QY 301 LLVGFRVAPARDKSKWVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTTWASWPIITILVLSV 360  
Db 301 LLVGFRVAPARDKSKWVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTTWASWPIITILVLSV 360  
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTPNRSY 420  
Db 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTPNRSY 420  
QY 421 RYDSLLLGKPNFSGILDLDLLELELERLHLQVWSPQAQRNLSLODICYAPLNPDNT 480  
Db 421 RYDSLLLGKPNFSGILDLDLLELELERLHLQVWSPQAQRNLSLODICYAPLNPDNT 480  
QY 481 SLYDCCINSLLQYFONNETLALLTANOTLMGOTSQVDKDHPLYCANAPLTIKDGTALAL 540  
Db 481 SLYDCCINSLLQYFONNETLALLTANOTLMGOTSQVDKDHPLYCANAPLTIKDGTALAL 540  
QY 541 SCMADYGAPVPFFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLOAKLWEAEAFLEEM 600  
Db 541 SCMADYGAPVPFFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLOAKLWEAEAFLEEM 600  
QY 601 RAFQRMAGMGMQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSWSRV 660  
Db 601 RAFQRMAGMGMQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSWSRV 660  
QY 661 MVDSKATLGLGGVAVVLGAVMAAGFFSYLGRSLVILQVVPFLVSVGADNIIFFVLE 720  
Db 661 MVDSKATLGLGGVAVVLGAVMAAGFFSYLGRSLVILQVVPFLVSVGADNIIFFVLE 720  
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSBAICFFLGALTMPMPAVRTFALTSLAV 780  
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSBAICFFLGALTMPMPAVRTFALTSLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVCKPQELPPPGQEGGLLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVCKPQELPPPGQEGGLLGFQKAYAPFL 840  
QY 841 LHWITRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYILDYFLFLNRYFEVGPV 900  
Db 841 LHWITRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYILDYFLFLNRYFEVGPV 900  
QY 901 YFVVTLLGYNFSEAGMAICSSACGNFSPFKIQYATERPEQSYLAIPASSWVDDFTDW 960  
Db 901 YFVVTLLGYNFSEAGMAICSSACGNFSPFKIQYATERPEQSYLAIPASSWVDDFTDW 960  
QY 961 LTPSSCRLYISGPNKDKFCPSTVNSLNCNLCNMSITMGSVRPSVEQFKHLYPFLNDRP 1020  
Db 961 LTPSSCRLYISGPNKDKFCPSTVNSLNCNLCNMSITMGSVRPSVEQFKHLYPFLNDRP 1020  
QY 1021 NIKCPKGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053  
Db 1021 NIKCPKGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053



QY 901 YFVTTTLYNFSSEAGNAICSSAGCANNFSTQKIQVATEPPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTTLYNFSSEAGNAICSSAGCANNFSTQKIQVATEPPEQSYLAIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVQFHKYLPWFNLNDP 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVQFHKYLPWFNLNDP 1020  
QY 1021 NIKCPKGLAAYSTVNLTSDDGQVL-----ASRPMAYH 1053  
Db 1021 NIKCPKGLAAYSTVNLTSDDGQVLDVVALSPRLEYSGTISAHCNLYLLDLSASRPMAYH 1080  
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPDTPDAFEVPPYITTNVFEQYLTILPEGL 1113  
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPDTPDAFEVPPYITTNVFEQYLTILPEGL 1140  
QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSIYVILVDTVGFNALWDISYNAVSLNLVVS 1173  
Db 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSIYVILVDTVGFNALWGISYNAVSLNLVVS 1200  
QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260  
QY 1234 LIQIFPFRNLNLTLLGLLHGLVFLPVLVSYGPDVNPALALEQKRAEBAVAAMVVASCP 1293  
Db 1261 LIQIFPFRNLNLTLLGLLHGLVFLPVLVSYGPDVNPALALEQKRAEBAVAAMVVASCP 1320  
QY 1294 NHPSRVSTADNIYVNHSPFGSIGKAGCAISNFLPNNGROF 1332  
Db 1321 NHPSRVSTADNIYVNHSPFGSIGKAGCAISNFLPNNGROF 1359

RESULT 8

ID ADY60780 standard; protein; 1359 AA.  
XX AC ADY60780;  
XX DT 19-MAY-2005 (first entry)  
XX DE Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 44.  
XX KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;  
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;  
KW metabolic disorder; atherosclerosis; cardiovascular disease;  
KW coronary artery disease; cerebrovascular ischemia; neurological disease;  
KW arteriosclerosis; transgenic animal.  
XX OS Homo sapiens.  
XX PN WO2005015988-A1.  
XX XX 24-FEB-2005.  
XX PF 16-DEC-2003; 2003WO-US040113.  
XX PR 17-JUL-2003; 2003US-00621758.  
PR 22-AUG-2003; 2003US-00646301.  
PR 16-SEP-2003; 2003US-00663208.  
XX PA (SCHE ) SCHERING CORP.  
XX PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;  
XX WPI; 2005-284403/29.  
DR N-PSDB; ADY60779.  
XX PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of  
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for  
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.  
XX

PS The invention relates to an isolated Niemann-Pick C1-like protein 1  
XX (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an  
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,  
CC human and mouse NPC1L1). Also included are an isolated polynucleotide  
CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a  
CC recombinant vector comprising the polynucleotide, a host cell comprising  
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or  
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),  
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous  
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not  
CC produce any functional NPC1L1 protein), an offspring or progeny of the  
CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a  
CC sample for an intestinal sterol or 5alpha-sterol absorption antagonist,  
CC inhibiting NPC1L1 mediated sterol or 5alpha-sterol uptake in a subject, a  
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical  
CC dosage form, and information indicating that NPC1L1 is a target of  
CC ezetimibe or substituted azetidinone), decreasing the level of intestinal  
CC sterol or 5alpha-sterol absorption in a subject (involving reducing the  
CC level of expression of NPC1L1 in the subject), identifying an antagonist  
CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a  
CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist  
CC of NPC1L1 which is useful for inhibiting or decreasing the level of  
CC NPC1L1 mediated sterol or 5alpha-sterol uptake in a subject such as  
CC mouse, rat or human. The level of expression of NPC1L1 in the subject is  
CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is  
CC useful for screening a sample for intestinal sterol or 5alpha-sterol  
CC absorption antagonist. The NPC1L1 antagonists identified are useful for  
CC treating medical conditions caused or mediated by NPC1L1, e.g.,  
CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or  
CC arteriosclerosis. The present sequence represents a mammalian NPC1L1  
XX protein.  
SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;  
QY 1 MAEAGIRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKPELSGSLMTLSNVCLSN 60  
Db 1 MAEAGIRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKPELSGSLMTLSNVCLSN 60  
QY 61 TPARKITGDHLLILQKICPRLVTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILQKICPRLVTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120  
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVPAA 180  
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVPAA 180  
QY 181 ATLAVGTWCVYGSALCNAQRWLNFGDTGNGLAFLDITFHLLEPGQAVSGIQPLNEGV 240  
Db 181 ATLAVGTWCVYGSALCNAQRWLNFGDTGNGLAFLDITFHLLEPGQAVSGIQPLNEGV 240  
QY 241 ARCNESQGDVATCSCQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300  
Db 241 ARCNESQGDVATCSCQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300  
QY 301 LLVGFVRVAPARDKSRWVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360  
Db 301 LLVGFVRVAPARDKSRWVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360  
QY 361 IPVVALAAGLVPTETLTPDVELWSAPNSQARSEKAFHQHFGPPFRTRNQVILTAPNRSSY 420  
Db 361 IPVVALAAGLVPTETLTPDVELWSAPNSQARSEKAFHQHFGPPFRTRNQVILTAPNRSSY 420  
QY 421 RYDSLILGPKNFSGLDLDLLELELELERLHLOVWSPQAQRNLSLODICYAPLNPDNT 480  
Db 421 RYDSLILGPKNFSGLDLDLLELELELERLHLOVWSPQAQRNLSLODICYAPLNPDNT 480  
QY 481 SLYDCINSLAQYFQNNRTLLLTANQTLMGQTSQVDWKDHPFLYCANAPLTPKDGATLAL 540

Db 481 SLVDCINSLLQYFQNNRTLLLTANQTLMGQTSQVDDKDFLYCANAPLTFKDGDTALAL 540  
Qy 541 SCADYCAPVFPFLAIGGYGKDYSEAEALIMTFSLNNYPAGDPRLAQAQKLWEAEFLEEM 600  
Db 541 SCADYCAPVFPFLAIGGYGKDYSEAEALIMTFSLNNYPAGDPRLAQAQKLWEAEFLEEM 600  
Qy 601 RAFORMAGMFQVTFTEAERSLEDEINRTTAEDLPFATPSYIVIFLYISLALGSYSNSRV 660  
Db 601 RAFORMAGMFQVTFTEAERSLEDEINRTTAEDLPFATPSYIVIFLYISLALGSYSNSRV 660  
Qy 661 MVDKSKATLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDKSKATLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720  
Qy 721 YQRLPRPGPREVHI GRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVTFEALTSGIAY 780  
Db 721 YQRLPRPGPREVHI GRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVTFEALTSGIAY 780  
Qy 781 ILDFLLQMSAFVALLSDSKRQASRLDCCVKPQELPPPGQEGELLLGFFOKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSDSKRQASRLDCCVKPQELPPPGQEGELLLGFFOKAYAPFL 840  
Qy 841 LHMTRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHMTRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Qy 901 YFVTTLYNFSSEAGMAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWDDFIDW 960  
Db 901 YFVTTLYNFSSEAGMAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWDDFIDW 960  
Qy 961 LTPSSCCRLYISGPNKDKFCPTSVNSLNCNKMSITWGSVRPVEQFHKYLPWFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPTSVNSLNCNKMSITWGSVRPVEQFHKYLPWFLNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVL-----ASRFWAYH 1053  
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLDTVAILSPRLEYSGTISAHCNLYLLDSASRFWAYH 1080  
Qy 1054 KPLKNSQDYETALRAARELANITADLRKVPDTPAPEVPPYTTTNVYFEQYLTIPEGL 1113  
Db 1081 KPLKNSQDYETALRAARELANITADLRKVPDTPAPEVPPYTTTNVYFEQYLTIPEGL 1140  
Qy 1114 FMLSICLVPTFAVSCLLGLDLRSLNLISIVMLVDTVGFMALWDISYNAVSLINLVS 1173  
Db 1141 FMLSICLVPTFAVSCLLGLDLRSLNLISIVMLVDTVGFMALWGISYNAVSLINLVS 1200  
Qy 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260  
Qy 1234 LIQIFPFRNLILITLLGLLHGLVFLPVILSVGDPVNPALALEOKRAEEA VAAVMVASCP 1293  
Db 1261 LIQIFPFRNLILITLLGLLHGLVFLPVILSVGDPVNPALALEOKRAEEA VAAVMVASCP 1320  
Qy 1294 NHPRSVSTADNIYNHNSPEGSIKGAGATSNFLPNNRGF 1332  
Db 1321 NHPRSVSTADNIYNHNSPEGSIKGAGATSNFLPNNRGF 1359

RESULT 9

AE93610 standard; protein; 1359 AA.

XX AC AE93610;

XX DT 06-OCT-2005 (first entry)

XX DE Human NPC1L1 protein.

XX KW Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;  
KW KD cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;  
KW KW arteriosclerosis; ds.

XX OS Homo sapiens.  
XX PN WO2005069900-A2.  
XX PD 04-AUG-2005.  
XX PF 14-JAN-2005; 2005WO-US001469.  
XX PR 16-JAN-2004; 2004US-0537341P.  
XX PA (MERI ) MERCK & CO INC.  
XX PI Garcia-Calvo M;  
XX PS WPI; 2005-564070/57.  
XX DR N-PSDB; AEB93609.  
XX DR GENBANK; AF192522.  
XX PT Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by contacting NPC1L1 with detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining if compound binds to human NPC1L1.  
XX PS Disclosure; SEQ ID NO 44; 215pp; English.  
XX CC The invention relates to identifying a ligand of NPC1L1. The method involves contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC1L1. In identifying a ligand of NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands are useful for stimulating or blocking the activity of NPC1L1, and for treating conditions caused or mediated by NPC1L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents the amino acid sequence of human NPC1L1 as disclosed under GenBank accession number AF192522.  
XX SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRCWLLWALLRLAQSEPYTTTHQPGYCAFYDECCGNPELSGLMTLSNVCLSN 60  
Db 1 MAEAGLRCWLLWALLRLAQSEPYTTTHQPGYCAFYDECCGNPELSGLMTLSNVCLSN 60  
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120  
Qy 121 VNLHCHNTCSNQSFLFNTRVAQLGAGQLPVAVYEAFTYOHSPAEOYSYDCSRRVPAA 180  
Db 121 VNLHCHNTCSNQSFLFNTRVAQLGAGQLPVAVYEAFTYOHSPAEOYSYDCSRRVPAA 180  
Qy 181 ATLAVGTCMGYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGAVGSGIQLNEGV 240  
Db 181 ATLAVGTCMGYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGAVGSGIQLNEGV 240  
Qy 241 ARCNESQGDVATCSQDCAASCPAIARPOLADSTFYLGQMPGSLVLIILLCSVFAVVTI 300  
Db 241 ARCNESQGDVATCSQDCAASCPAIARPOLADSTFYLGQMPGSLVLIILLCSVFAVVTI 300  
Qy 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPQGGTGWVWASWPLTILVSV 360  
Db 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPQGGTGWVWASWPLTILVSV 360  
Qy 361 IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGPFRTNQVILTAPNRSY 420  
Db 361 IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGPFRTNQVILTAPNRSY 420

QY	421	RYDSLLGPKNFGSILDLDLLLELLELQERLRLQVWSPQAQNI	SLQDICTYAPLNP	480
Db	421	RYDSLLGPKNFGSILDLDLLLELLELQERLRLQVWSPQAQNI	SLQDICTYAPLNP	480
QY	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHF	LYCANAPLTPKDG	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHF	LYCANAPLTPKDG	540
QY	541	SCMADYGAPVFPFLAIGGYKGKDYSEAEALIMTFFSLNYP	PAGDPRLAQAKLWEEAFLEEM	600
Db	541	SCMADYGAPVFPFLAIGGYKGKDYSEAEALIMTFFSLNYP	PAGDPRLAQAKLWEEAFLEEM	600
QY	601	RAFQRMAGHFQVTFTAERLSLEDEINRTTAEDLP	IPFATSYIVIFLYISLALGSYSWSRV	660
Db	601	RAFQRMAGHFQVTFMAERLSLEDEINRTTAEDLP	IPFATSYIVIFLYISLALGSYSWSRV	660
QY	661	MVDSKATLGLGGVAVVLGVAAMAGPSPSYLGIBSSVIL	QVVPFLVLSGADNIPFVLE	720
Db	661	MVDSKATLGLGGVAVVLGVAAMAGPSPSYLGIBSSVIL	QVVPFLVLSGADNIPFVLE	720
QY	721	YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEAL	ICFFLGALTPMPAVRTPALTSGLAV	780
Db	721	YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEAL	ICFFLGALTPMPAVRTPALTSGLAV	780
QY	781	ILDFLQMSAFVALLSLDSKROBASRLDVCCCKVQ	PQELPPPGQEGLLLGFFQKAYAPFL	840
Db	781	ILDFLQMSAFVALLSLDSKROBASRLDVCCCKVQ	PQELPPPGQEGLLLGFFQKAYAPFL	840
QY	841	LHWITRGVULLFLALPGVSLYSMCHISVGLDQELAL	PKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVULLFLALPGVSLYSMCHISVGLDQELAL	PKDSYLLDYFLFLNRYFEVGAPV	900
QY	901	YFVTTLGYNFSSAGNAICSSAGCANNFSTQIKQI	QVATFPPEOSYLAIPASSWVDDFIDW	960
Db	901	YFVTTLGYNFSSAGNAICSSAGCANNFSTQIKQI	QVATFPPEOSYLAIPASSWVDDFIDW	960
QY	961	LTPSSCCRLYISGPNKDKFCPTVNSLNC	LKNCMSITMGSVRPSVQPHKYL	1020
Db	961	LTPSSCCRLYISGPNKDKFCPTVNSLNC	LKNCMSITMGSVRPSVQPHKYL	1020
QY	1021	NIKCPKGGIAAYSTSVNLTSDGQVL-----	ASRPMAYH	1053
Db	1021	NIKCPKGGIAAYSTSVNLTSDGQVL-----	ASRPMAYH	1053
QY	1054	KPLKNSQDYTEALRAARELAAMITADLRKVP	GTDPAPFVPPYTTITNVFEQYL	1113
Db	1081	KPLKNSQDYTEALRAARELAAMITADLRKVP	GTDPAPFVPPYTTITNVFEQYL	1140
QY	1114	FMLSCLVPTFAVSCLLGLDRLSGLNLSIVML	LDVTVGFMALWDISYNVSLINLVS	1173
Db	1141	FMLSCLVPTFAVSCLLGLDRLSGLNLSIVML	LDVTVGFMALWDISYNVSLINLVS	1200
QY	1174	AVGMSVEFVSHITRSPAISTKTWLERAKEATIS	MGSAVAGVAMTNLPGLI	1233
Db	1201	AVGMSVEFVSHITRSPAISTKTWLERAKEATIS	MGSAVAGVAMTNLPGLI	1260
QY	1234	LQIQIFFFRLNLITLLGLHGLVFLPVLISY	YVGPVNPALAEOKRAEAAVAVASCP	1293
Db	1261	LQIQIFFFRLNLITLLGLHGLVFLPVLISY	YVGPVNPALAEOKRAEAAVAVASCP	1320
QY	1294	NHPSRVSTADNIYVNHSPFGSISKGAGAI	SNFLPNNGRQF	1332
Db	1321	NHPSRVSTADNIYVNHSPFGSISKGAGAI	SNFLPNNGRQF	1359
RESULT 10				
ID	ABG22693	standard; protein; 1344 AA.		
XX	ABG22693			
AC	ABG22693			
XX	18-FEB-2002	(first entry)		
DT				

XX	Novel human diagnostic protein #22684.			
DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;			
XX	food supplement; medical imaging; diagnostic; genetic disorder.			
KW				
XX	Homo sapiens.			
OS				
XX	WO200175067-A2.			
PN	11-OCT-2001.			
XX				
PD	30-MAR-2001; 2001WO-US008631.			
XX				
PF	31-MAR-2000; 2000US-00540217.			
XX				
PR	23-AUG-2000; 2000US-00649167.			
XX	(HYSE-) HYSEQ INC.			
PA	Drmanac RT, Liu C, Tang YT;			
XX	WPI; 2001-639362/73.			
XX	N-PSDB; AAS86880.			
DR	New isolated polynucleotide and encoded polypeptides, useful in			
XX	diagnostics, forensics, gene mapping, identification of mutations			
PT	responsible for genetic disorders or other traits and to assess			
PT	biodiversity.			
XX	Claim 20; SEQ ID NO 53052; 103pp; English.			
PS	The invention relates to isolated polynucleotide (I) and polypeptide (II)			
XX	sequences. (I) is useful as hybridisation probes, polymerase chain			
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,			
CC	and in recombinant production of (II). The polynucleotides are also used			
CC	in diagnostics as expressed sequence tags for identifying expressed			
CC	genes. (I) is useful in gene therapy techniques to restore normal			
CC	activity of (II) or to treat disease states involving (II). (II) is			
CC	useful for generating antibodies against it, detecting or quantitating a			
CC	polypeptide in tissue, as molecular weight markers and as a food			
CC	supplement. (II) and its binding partners are useful in medical imaging			
CC	of sites expressing (II). (I) and (II) are useful for treating disorders			
CC	involving aberrant protein expression or biological activity. The			
CC	polypeptide and polynucleotide sequences have applications in			
CC	diagnostics, forensics, gene mapping, identification of mutations			
CC	responsible for genetic disorders or other traits to assess biodiversity			
CC	and to produce other types of data and products dependent on DNA and			
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic			
CC	patent did not appear in the invention. Note: The sequence data for this			
CC	electronic format directly from WIPO at			
CC	ftp.wipo.int/pub/published_pct_sequences			
XX	Sequence 1344 AA;			
SQ				
Query Match 94.6%; Score 6536; DB 4; Length 1344;				
Best Local Similarity 96.7%; Pred. No. 0;				
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;				
QY	1	MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNV	SCLSN	60
Db	1	MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNV	SCLSN	60
QY	61	TPARKITGDHLLTLLOKICPRLYTGNTQACSAKOLVSLASLSITKALLTRCP	ACSDNF	120
Db	61	TPARKITGDHLLTLLOKICPRLYTGNTQACSAKOLVSLASLSITKALLTRCP	ACSDNF	120
QY	121	VNLHCHNTCSPNOSLFINVTRVAQAGAGQLPAVVAEAFYQHSFAEQSYDSCSR	VRVPA	180
Db	121	VNLHCHNTCSPNOSLFINVTRVAQAGAGQLPAVVAEAFYQHSFAEQSYDSCSR	VRVPA	180
QY	181	ATLAVGTMCVGVSGALCNAQRWLNFGDGTGNGLAPLDTITFHLRPGQAVSG	IQIPLNEGV	240



Db 60 TPAHVHTGEHALLQRICPLRYNGPNTTFACCSTKQLLSLESSMSITKALLTRCPACSDN 119  
Qy 120 FVNHLCHNTCSNPQSLFINVTRVAQAGAGQLPAVAYEAFYQHSFAEQSDYSCSRVRVPA 179  
Db 120 FVSLHCHNTCSNPQSLFINVTRVVERGAGEPPAVAYEAFYQSFPAEKAYESCQVRIPA 179  
Qy 180 AATLVACTMCGVYGSALCNARWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQLNEG 239  
Db 180 AASLVAGSMCGVYGSALCNARWLNFGQDGTGNGLAPLDITFHLLEPGQALPDGIQLNGK 239  
Qy 240 VARCNESQGDVDTCSQDCAASCAPAIARPOALDSTFYLGMQPGSLVLIILGCSVPAVVT 299  
Db 240 IAPCNESQGDSDSVCQDCAASCVPVPPPEALRPSPFYMGMFGWGLALIIIFTVAVFVLS 299  
Qy 300 ILLVGFVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFPQCGMTWVASWPLTILVLS 359  
Db 300 AVLVLRLVSNRNKNAEGPQAPKLPKHKLSPHTILGRFQNGWGRTRVASWPLTVLALS 359  
Qy 360 VIPVVALAAGLVETELTDPVELWSAPNSQARSEKAFHDOHFGPPFRFTNQVILTAPNRSS 419  
Db 360 FIVVIALAAGLTFIELTDPVELWSAPKSOARKEKSFHDEHFGPFRTNQIFVTARNRSS 419  
Qy 420 YRYDSLILGPNFSGILDLDLLELELERLRLHQLWSPEAORNISLQDICVAPLNPDN 479  
Db 420 YKYSLSLLGSKNFSGLISLDFLELELERLRLHQLWSPEAERNISLQDICVAPLNPN 479  
Qy 480 TSLYDCCINSLLQYFQNNRTLLLTANTQTLGQTSQVDMKDHFLYCANAPLTFKDGITALA 539  
Db 480 TSLSDCCVNSLLQYFQNNRTLLMTANTQTLGQTSVDMKDHFLYCANAPLTFKDGTSLA 539  
Qy 540 LSCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNPNYPAGDPRLAQAKWEEAFLE 599  
Db 540 LSCMADYGAPVFPFLAVGGYQGYDYSEAEALITFSLNPNYPAGDPRMAQAKWEEAFLE 599  
Qy 600 MRAFORMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVFLYISLALGYSWSR 659  
Db 600 MESFQRTSDKQVAFSAERKSEDEINRTTIQDLPVFAVSIIIVFLYISLALGYSRCRS 659  
Qy 660 VMVDSKATLGLGVAVVLGAVMAAGFFSYLIGRISLIVLIQVVPFLVLSVGADNIFVL 719  
Db 660 VAVESKATLGLGVIVVLGAVLAAGFYSLGVPSLIVLIQVVPFLVLAAGDNIIFVL 719  
Qy 720 EYQRLPRRPGEPREHIGRALGVAPSMLLCSLSEAI CFFPLGALTTPMPAVRTPALTSGLA 779  
Db 720 EYQRLPRMPGEQREAHIGRTLGSVAPSMLLCSLSEAI CFFPLGALTTPMPAVRTPALTSGLA 779  
Qy 780 VILDFTLQMSAFVALLSLDSKQESASRLDVCCCKPQELPPPGQGGLLGFFQKAYAPF 839  
Db 780 IILDFTLQMTAFVALLSLDSKQESASRPDLVCCFSTRKUPPPPKKESGLLRFFRKIYAPF 839  
Qy 840 LLHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVVGAP 899  
Db 840 LLHRFTRPVVMLLFLTLFGANLYLMCNINVLGDQELALPKDSYLLIDYFLFLNRYLEVGP 899  
Qy 900 VIFVTTLGNFSEAGNNAICSSAGCNFSFTQKIQYATEFPQSYLAIPASSWVDDFD 959  
Db 900 VIFVTTSFGNFSEAGNNAICSSAGCKSFSLTQKIQYASEFPQSYVAIAASSWVDDFD 959  
Qy 960 WLTP-SSCCRLYISGNKDKFCBSTVNSLNCNLSITWGSVRPSVQFHKYLPFLND 1018  
Db 960 WLTPSSCCRLYIRGHKDEFCBSTDTSFNCLKNCNRTLGPVRPTABQFHKYLPFLND 1019  
Qy 1019 RPNKCPKGLAAYSTSVNLTDQVLAASFMAHYKPLKNSDYETALAAARELANITA 1078  
Db 1020 PPNIRCPKGLAAYRTSVNLSDDQVIAQFMAHYKPLNSQDFTTALASRLAANITA 1079  
Qy 1079 DLKRVGTDPAFEPFVYITNVFEYQYLTLPGLFMLSLCLVPTFAVSCLLGLDLRG 1138  
Db 1080 DLKRVGTDPEFVFPYITSNVFEYQYLTLPGLFMLSLCLVPTFAVSCVLLGLDMCSG 1139  
Qy 1139 LLNLLSIIMLVDTVCFMALWDISYNAVSLINLVSAGVSEFVSHITSPALSTKPTWL 1198  
Db 1140 ILNLLSIIMLVDTIGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITSPAVSTKPTRL 1199

Qy 1199 ERAKEATISGSAVPAGVAMTNLPGLVLGLAKAQIQLIFFRNLITLLGLLHGLVFL 1258  
Db 1200 ERAKDATVFMGSAVPAGVAMTNFPFGLILGLFAQAQLIQLIFFRNLITLLGLLHGLVFL 1259  
Qy 1259 PVTLISYVGDVNPALALEQKBAEAAVAVMVASCPNHPSRVSTADNIYVNHSEGSIGK 1317  
Db 1260 FVLSYLGVDVNLQALVQEEKLSEA-AVAPEPCQPYPSPADADAN--VNYGFAPELAHG 1316  
Qy 1318 AGAISNLFNNGRQF 1332  
Db 1317 ANAARSLPKSDQKF 1331

## RESULT 12

ID ADY60738  
ADY60738 standard; protein; 1331 AA.

XX ADY60738;

XX DT 19-MAY-2005 (first entry)

XX Rat Niemann-Pick C1-like protein 1 (NPC1L1) protein.

XX Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;  
KW Cardiant; Vasotropic; Cerebroprotective; Cholesterol; hyperlipidemia;  
KW metabolic disorder; atherosclerosis; cardiovascular disease;  
KW coronary artery disease; cerebrovascular ischemia; neurological disease;  
KW arteriosclerosis; transgenic animal.

XX Rattus sp.

OS WO2005015988-A1.

XX 24-FEB-2005.

XX 16-DEC-2003; 2003WO-US040113.

XX 17-JUL-2003; 2003US-00621758.

XX 22-AUG-2003; 2003US-00646301.

XX 16-SEP-2003; 2003US-00663208.

XX (SCHE ) SCHERING CORP.

XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX WPI: 2005-284403/29.

XX N-PSDB; ADY60737, ADY60746.

PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of

PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for

PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.

XX Claim 30; SEQ ID NO 2; 146pp; English.

XX The invention relates to an isolated Niemann-Pick C1-like protein 1  
(NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an  
amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,  
human and Mouse NPC1L1). Also included are an isolated polynucleotide  
encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a  
recombinant vector comprising the polynucleotide, a host cell comprising  
the vector, an isolated antibody which specifically binds to NPC1L1 (or  
to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778).

XX preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous  
mutation of endogenous chromosomal NPC1L1 (where the mouse does not  
produce any functional NPC1L1 protein), an offspring or progeny of the  
mouse having inherited a mutated NPC1L1 allele of the mouse, screening a  
sample for an intestinal sterol or Salpha-sterol absorption antagonist,  
inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a  
kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical  
dosage form, and information indicating that NPC1L1 is a target of  
ezetimibe or substituted azetidinone), decreasing the level of intestinal  
sterol or Salpha-sterol absorption in a subject (involving reducing the



PT glucuronide and a candidate compound and determining if compound binds to  
XX human NPC1L1.

Example 2; SEQ ID NO 2; 215pp; English.

CC The invention relates to identifying a ligand of NPC1L1. The method  
CC involves contacting human NPC1L1 with a detectably labeled substituted 2-  
CC azetidinone glucuronide and a candidate compound and determining whether  
CC the candidate compound binds to human NPC1L1. In identifying a ligand of  
CC NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10  
CC nM or lower. The detectably labeled substituted 2-azetidinone glucuronide  
CC is a 35 S-labeled compound 2, given in the specification. NPC1L1 ligands  
CC are useful for stimulating or blocking the activity of NPC1L1, and for  
CC treating conditions caused or mediated by NPC1L1. It is useful for  
CC reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart  
CC disease, stroke, or arteriosclerosis. The present sequence represents a  
CC rat NPC1L1, a N-glycosylated protein.

XX Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;  
Best Local Similarity 77.8%; Pred. NO. 0;  
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAAAGLRLGMLLWALLRLAQSEPYTHHOPGYCAFDCECKNPGLSGSLMTLSNVSCLSN 60  
DB 1 MAAAWL-GMLLWALLLSAAQGLYTPKHEAGVCTFVECKNPGLSGGLTSLNSVSCLSN 59  
QY 61 TPARKITGHLILLOKICRLPYGNPTQ-ACCSAKOLVLSLEASLITKALLTRCPACSDN 119  
DB 60 TPARKITGHLILLOKICRLPYGNPTQ-ACCSAKOLVLSLEASLITKALLTRCPACSDN 119  
QY 120 FVNLHCHNTCSNPQSLFINVTRVAOLGAGOLPAVVAEYAFYQSHFAEQSDSCSRVRVPA 179  
DB 120 FVSLHCHNTCSNPQSLFINVTRVVERGAGEPPAVVAEYAFYQSHFAEQSDSCSRVRVPA 179  
QY 180 AATLAVGTWCVGYGSLACNAQRWLNFGDGTGNGLAPLDTFHLLEPGQAVGSGIOPING 239  
DB 180 AASLAVGSCVGYGSLACNAQRWLNFGDGTGNGLAPLDTFHLLEPGQAVGSGIOPING 239  
QY 240 VARNESQGDVATCSQDCQCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVT 299  
DB 240 IAPNESQGDVATCSQDCQCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVT 299  
QY 300 ILLVGRVAPARDKSRVDPKKGTSLSKLSFSTHTLLGQFQCGMTWVASWPLTLVL 359  
DB 300 AVLVRLRVSVNRNKNKAEGQEAQPKLPHKHLSPHITLGRFFQNGWTRVASWPLTLVL 359  
QY 360 VIPVVALAAGLVPTTEITDVELWSPNSOARSEKAFHDOHGFPPFRITNOVILTAPNRS 419  
DB 360 FIVVIALAAGLVPTTEITDVELWSPNSOARSEKAFHDOHGFPPFRITNOVILTAPNRS 419  
QY 420 YRYSLLGPKNFGSLDLDLLELELQERLHQLVWSPQARNTSLQDICVAPLNPN 479  
DB 420 YRYSLLGPKNFGSLDLDLLELELQERLHQLVWSPQARNTSLQDICVAPLNPN 479  
QY 480 TSLYDCCINSLQYFONNRTLLLTANQTLMGTSQVDKDFLYCANAPLTFKDTALA 539  
DB 480 TSLSDCCVNSLQYFONNRTLLMTANQTLMGTSQVDKDFLYCANAPLTFKDTALA 539  
QY 540 LSCMADYGAVPFFFLAIGYKGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 599  
DB 540 LSCMADYGAVPFFFLAIGYKGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 599  
QY 600 MEAFQRMAGMQVTTAERSLEDEINRTAEDLPITFATSYIVIFLYISLALGSYSSWR 659  
DB 600 MESFQRTSDKQVAFSAERSLEDEINRTIQDLPVFAVSYIVIFLYISLALGSYSSWR 659  
QY 660 VMVDSKATLGLGVAVVLGAVMAMGFYSYGLRSLVLQVVPFVLVSGADNIFIVL 719  
DB 660 VAVESKATLGLGVAVVLGAVMAMGFYSYGLRSLVLQVVPFVLVSGADNIFIVL 719  
QY 720 EYQRLPRRPGEPREHIGRALGRVAFPSMLLSLSEAIKCFPLGALTMPAVRTFALTSGLA 779

DB 720 EYQRLPRRPGEPREHIGRALGRVAFPSMLLSLSEAIKCFPLGALTMPAVRTFALTSGLA 779  
QY 780 VILDDELLQMSAFVALLSLDSKQESRLDVCCKVKQDELPPPGQGGGLLGFQKAYAPF 839  
DB 780 IILDDELLQMTAFVALLSLDSKQESRPDLVCCFSTRKLLPPPKKESGLLRLFRKTYAPF 839  
QY 840 LLHWITRGVVLALLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAP 899  
DB 840 LLHRTIRPVVMLLFTLFGANLYLMCNINVGLDQELALPKDSYLLDYFLFLNRYLEVGP 899  
QY 900 VYFVITLGNVPSSEAGMNAICSSAGCNFSFTQIOYATEFFEQSYLAIPASSWDDFID 959  
DB 900 VYFVITSGFNFSSEAGMNAICSSAGCKSFLTKIOYASEFFPDQSYVAIAASSWDDFID 959  
QY 960 WLTP-SSCCRLYISGNKDKFCPSTVNSLNCNCKNSITMGSVRPSVEQFHKYLWFLND 1018  
DB 960 WLTPSSCCRLYIRGPHKDFCPSTDTSFNCLNCKNRTLGPVRPFAEQFHKYLWFLND 1019  
QY 1019 RPNKCPKGGLAAYSTSVNLTSDGQVLSRFRMAYHKPLKNSODYTEALRAARELANITA 1078  
DB 1020 PNIIRCPKGGLAAYRTSVNLTSDGQVLSRFRMAYHKPLKNSQDFTALRASRLAANITA 1079  
QY 1079 DLKVPGTDPATEVFPYTTITNVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRS 1138  
DB 1080 DLKVPGTDPNFEVFPYTTISNVFYQOYLTVLPEGIPTLALCFVPTFVVCVLLGLDWC 1139  
QY 1139 LNLLSIVMLVDTVGFMAWLDISYNAVSLINLSAVGMSVFSVSHITSEFAISTKPTWL 1198  
DB 1140 IINLSIIMLVDTITGLMAVWGISYNAVSLINLTVAGMSVFSVSHITSEFAISTKPTWL 1199  
QY 1199 ERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQILQIFFRLNLLITLLGLLHGLVFL 1258  
DB 1200 ERAKDATVMSGSAVAGVAMTNFPGILLGLGFAQALQILQIFFRLNLLITLLGLLHGLVFL 1259  
QY 1259 PVILSVGPDVNPALALEOKRAEAAVAVMVASCNHPRSVSTADNINYHNSPEGI-KG 1317  
DB 1260 PVVLSYLGPDVNPQALVQEEKLASEA-AVAPEPCPOYPSPADADAN--VNYGPAPELAHG 1316  
QY 1318 AGAISNFIPLNNGROF 1332  
DB 1317 ANAARSSLPKSDQKF 1331  
RESULT 14  
ADJ27305  
ID ADJ27305 standard; protein; 1333 AA.  
XX AC ADJ27305;  
XX DT 20-MAY-2004 (first entry)  
XX DE Mouse NPC1L1.  
XX KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;  
KW plasma membrane; transport signal; promoter;  
KW sterol regulated element binding protein 1; SREBP1;  
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;  
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;  
KW cholesterol absorption; serum cholesterol; hyperlipidemia;  
XX atherosclerosis; coronary heart disease; stroke; arteriosclerosis.  
OS Mus sp.  
XX WO2004009772-A2.  
XX 29-JAN-2004.  
XX 17-JUL-2003; 2003WO-US022467.  
XX 19-JUL-2002; 2002US-0397442P.  
XX (SCHE ) SCHERING CORP.  
PA







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:16:49 ; Search time 52 Seconds  
(without alignments)  
2464.629 Million cell updates/sec

Title: US-10-736-769-4  
Perfect score: 6909  
Sequence: 1 M8AGLGRGWLWALLRLAQ.....GSIKGAGAI5NLPNNGRQF 1332

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2385	34.5	1278	2 T30188	Niemann-Pick C dis
2	1394	20.2	1055	2 T05663	hypothetical prote
3	1347	19.5	1456	2 T15961	hypothetical prote
4	1329	19.2	1170	2 S55525	probable membrane
5	646.5	9.4	1434	2 T30172	transmembrane prot
6	628.5	9.1	1442	2 T18538	patched protein -
7	600	8.7	915	2 S44797	F09G8.4 protein -
8	598.5	8.7	933	2 T25600	hypothetical prote
9	583.5	8.4	1182	2 T13952	membrane protein p
10	573.5	8.3	1003	2 T26746	hypothetical prote
11	572	8.3	1220	2 T18291	patched protein -
12	539.5	7.8	889	2 T29590	hypothetical prote
13	514	7.4	1405	2 T27969	hypothetical prote
14	473.5	6.9	1015	2 T15830	hypothetical prote
15	464.5	6.7	1299	1 S06119	membrane protein p
16	458.5	6.6	955	2 T21612	hypothetical prote
17	436	6.3	956	2 A89153	protein C24B5.3 (i
18	409.5	5.9	936	2 T26521	hypothetical prote
19	407	5.9	800	2 T26683	hypothetical prote
20	399.5	5.8	881	2 T31739	hypothetical prote
21	381	5.5	840	2 T33217	hypothetical prote
22	379	5.5	714	2 T16126	hypothetical prote
23	367	5.3	690	2 T23399	hypothetical prote
24	367	5.3	877	2 T24097	hypothetical prote
25	349	5.1	845	2 T25657	hypothetical prote
26	347.5	5.0	820	2 T32908	hypothetical prote
27	312	4.5	890	2 T22186	hypothetical prote
28	297.5	4.3	871	2 T28706	hypothetical prote
29	294	4.3	983	2 T21213	hypothetical prote

30	276.5	4.0	633	2	S44795	F09G8.3 protein -
31	259.5	3.8	413	2	S28276	hypothetical prote
32	224	3.2	1276	2	T18526	SREBP cleavage act
33	189.5	2.7	1154	2	T48829	related to SREBP c
34	187.5	2.7	932	2	T28820	hypothetical prote
35	186.5	2.7	1227	2	T20370	hypothetical prote
36	179.5	2.6	787	2	H71453	hypothetical prote
37	173	2.5	969	2	T33156	hypothetical prote
38	170	2.5	746	2	A75018	transport protein
39	154	2.2	823	2	B81282	probable ABC-type
40	149	2.2	1011	2	T07712	probable ABC-type
41	149	2.2	1051	2	AG3455	multidrug resistan
42	148.5	2.1	746	2	H84301	hypothetical prote
43	146	2.1	1041	2	AC0423	multidrug efflux p
44	143	2.1	724	2	H69780	antibiotic transpo
45	141.5	2.0	1049	2	AF0561	acriflavin resista

ALIGNMENTS

RESULT 1

T30188  
Niemann-Pick C disease protein - mouse  
N;Alternate names: NCPI protein  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30188  
R;Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellis  
Science 277, 232-235, 1997  
A;Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis  
A;Reference number: Z20765; MUID:97362324; PMID:9211850  
A;Accession: T30188  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1278 <LOF>  
A;Cross-references: UNIPROT:Q35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:G2251241  
C;Genetics:  
A;Gene: Npc1  
A;Map position: 18

Query Match 34.5%; Score 2385; DB 2; Length 1278;  
Best Local Similarity 40.2%; Pred. No. 1.2e-158;  
Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY	14	LLRLAQSPYTTIHQPGYCAPVDECGKPELSGLMTLSNVSLNTSPARKITGDHLIL	73
DB	12	LLLLCPAQVFSQ-----SCVWIGCG---IATGD---KRYNCKYSGPKPLPKGYDL	59
QY	74	LQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHTCSNQ	133
DB	60	VQELCPGLFF-DNVSICCIDIQLQLKSNLQLPLQLSRCPSCFYNLMTLFCBLTCSPHQ	118
QY	134	SLFINVTRVAQIAGAGQLPA-----VVAYEAFYQHSFAEQYDSCSRVRVPAATLAVGTWC	189
DB	119	SOFLNVTATYEDFPKTPENTKNTVKELEYVQGFSANAMYNACRDVEAPSSSEKALGLLC	178
QY	190	GVYGSALCNAQWLNFGQDTGNGLAPLDI-----TFHLEPGQAVSGGIQLENEGVARN	244
DB	179	GDARA-CNATWIEYMFNKGNGQAPFTIIPVFSLSIL-----GMEPRNATKGCN	229
QY	245	ESQGDVATCSQDCAASC-----PAIARPOALDSTF-----YLGQMPGSLV	286
DB	230	ESVDEVTGSCQDCSIVCGPKQPPPPMPMPNRIWGLDAMYIMVYVAFVFPFGALL	289
QY	287	LIILCSFVAVVTILLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFGWT	346
DB	290	AVWCHRRRYFVSEYTPIDSNIAFVNS-----DKGEASCCDPLGAFAFDCLARMFTKGA	345
QY	347	WTASWPLTLVLVSPVVALAAGLVFTELTDPVLSAPNSQARSEKAFHQHGFPER	406
DB	346	FCVRNPTCIFFSLAFITVCSGSLVFVQVTINPVELWSAPHSQARLEKEYFKHGFPER	405



Db 770 -----:-----FAL 772  
QY 931 TQKIQVATEFPQSYLAIPASSWVDDFDLWLTSS--CCRLYISGPNKDKFCPSTVNSLN 988  
Db 773 ASIIISASQASDTSYAKPAASWLDLFLVWLSPEAFGCCCKFTNG-----SYCPP---DDQ 825  
QY 989 CLKNCMSITMGSVRPSVEQFHKYLPWFLNDRNRIKCPKGLAAYSTSNLT--SDQVLA 1046  
Db 826 CFRHS--DLVQDRPSTAQFREKLWFLNALPSADCAKGGHAYTNSVDLKGYESVIOA 882  
QY 1047 SFPMAYHKLKNSQDYTEALRAARELANITADLRKVPGTDPAFVFPYPTITNVFYEQL 1106  
Db 883 SBFRYHTPL-NTQ-----IDIFFYSVFVFPEQYL 912  
QY 1107 TILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSTVMILVDTVGFPMALWDISYNV 1166  
Db 913 NITWTALNTALAIV-----GIQLNAV 934  
QY 1167 SLINLVSAGMGEVFSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIV 1226  
Db 935 SVVNLIMSIGIAVEFCVHISHAFLMSSGDR-EHRAREALETWGAASFSGIITLKLGVIV 993  
QY 1227 LGLAKAQLIQIFFRNLNLTLLGLHGLVFLPVLISYVG-PDVNPALALEOKRAEAVA 1285  
Db 994 LCFARSEIPVYVYFQMYLALVIIGFLHGLVFLPVLISLAGPPQLN--LDIEQQQOTDEASS 1051  
QY 1286 AVN 1288  
Db 1052 SLI 1054

RESULT 3  
T15961  
hypotheical protein F02E8.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T15961  
R:Miller, N.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of C. elegans cosmid F02E8.  
A:Reference number: Z18436  
A:Accession: T15961  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1456 <MIL>  
A:Cross-references: UNIPROT:Q19127; UNIPARC:UPI000017B8AD; EMBL:U53340; NID:g1255859; PI  
A:Experimental source: strain Bristol N2; clone F02E8  
C:Genetics:  
A:Gene: CESP:F02E8.6  
A:Map position: X  
A:Introns: 71/1; 124/3; 218/3; 269/3; 310/3; 403/2; 484/3; 528/2; 764/1; 858/3; 888/3; 1

Query Match 19.5%; Score 1347; DB 2; Length 1456;  
Best Local Similarity 27.5%; Pred. No. 7.7e-86;  
Matches 372; Conservative 258; Mismatches 565; Indels 158; Gaps 41;  
QY 76 KICRLYTGPTQACCSAKOLVLEASLITKALTTRCPACSDNFVNLHCNTCSNQSL 135  
Db 63 EFCPHLLTGDN-KLCCTPSQAEGLTKQIAQARHILGRCPSCFDNFPAKLWCEFTCSNQD 121  
QY 136 FINTRVAQL--GAGQLPAVVAYEAF-----YQHS--PAEQSYDSCSRVRVPAAATLAVG 186  
Db 122 FVSISEMKPIKEKGGFTPEYQPAAYVNTVEYRSTDPAGNMFSCCKDVTGGQPALRV- 180  
QY 187 TMCVGYGALCAQRWLNFGDGTGNGL-APLIDITFHLLEPQAVGSGIQP-LNBSGARCN 244  
Db 181 -MC--TSTPCTLTNWLFEFIGTQNLDMNIPHTKFLLVDPKTPPSDRSTYMNVTGCD 236  
QY 245 ESQGDVATCSQDCAACPAIRPQALDSTFYLGOM-----PGSLVLIILICSVFAVVT 299  
Db 237 KSARVGWPACTSEC--NKEEYANLIDLDGKTSGQTCNVHGIACLNIFVNLAFIGSLAV 294

QY 300 ILLVGF-----RVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQPF 341  
Db 295 LLCVGFVTSYDEDTYNLRQTSGESPKENRIK-----RTGAWI-----INFENNA 342  
QY 342 QGWGTWVASWPLTILVLSVIPVVALAAGLVLTETTPVELWSAPNSQARSEKAFDQHF 401  
Db 343 RDIGMAGRNPKSHFPGICAVLIFCLPGMIYHKESTNVVDMSSPRSRARQEEMVFNANF 402  
QY 402 GPFFRTNQVILTAPNRSSVRYDSLGLPKNPFSGILDLLLELLELLEQLERHLHQVWSPA 461  
Db 403 GRPQRYQQIMLL--SHRDFQSSGKLYGP-----VFHKDIFEBELFDILNAIKNISTQSDG 455  
QY 462 QRNISLDIQCAYPLNPDNTSLYDCCINSLLQYFQNNRTLLLTANOT-----L 509  
Db 456 -RTIILDDVCYRPMGPG-----YDCLIMSPNTYFQGNKEHLMKSKKEETVEDDDADPXF 510  
QY 510 MGQTSQVDKWHKFLYCANAPLTFKDTALALSCMADYGAIPVFPFLAIGGYKGDYSEABA 569  
Db 511 SSEATTDEWMNHMAACIDQPMQK--TKSGLSCMGTYGGPSAPNM-VFCKNSTNHQAANS 567  
QY 570 LIMTFSLANNYPAGDPRLAQAQKLEEAFLIEMPAFORRMAGMFQVTFATRSLEDEINRT 629  
Db 568 IMMTILVTQ--RTEPIQKAELEWEKFLKFKKEYREKSPKVI-FSFMARERSITDEIENDA 624  
QY 630 AEDLPFATSYIVIFLYISLALGSY-----SSWSRVMYDVS KATLGLGGVAVVLGAVMAAM 684  
Db 625 KDEIVTVVIALAFLIGYVTFSLGRYPVCENQLWS-ILVHSRICLGLMWSIINLLSGFCSW 683  
QY 685 GFFSYLGRSSRLVLQVVPFLVLSVGADNIFIFVLEY--QRLP--RRPGEPREVHIGRA 739  
Db 684 GIFSMFGIHPKVALVQFVVVTLGVCRTFMVVKYQAQRVSMFVMSPDQCEI-VGNV 742  
QY 740 LGRVAPSMLLCSLSEAICFFLGAITPMPAVRTFALTSLGLAVLDLLOMSAFVALLSDS 799  
Db 743 MAGTMPAMFSSSLGCAPSFIFGGFTDLPAIRTFCLYAGLVLIDVVLHCTIFLALFVMDT 802  
QY 800 KQEBASRLDVCCKVPQELPPPQGEGL--LLG-----PFQKAVAPELLHWI 844  
Db 803 QRELNG-----KP-EFFPYQIKDLGLAYLIGRQRTDTFTWQFFHFQVAPFLMHM 853  
QY 845 TRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEYGAIFYVT 904  
Db 854 TRIITGIIFIASFITTVILSSKISVGPDSMAFTEKSYISTHFRYLDKDFDVGPPVFTV 913  
QY 905 TLGNFSEAGMNAICSSAGCNNSFTOKIOYATEFPEQSYLAIPASSWDDFDIWLTP 963  
Db 914 DGEIDWHRPDVQNKRFCTPPGCSDTSGFNIMNYAVGHTBQTYLSGEMYNWIDNLEWISRK 973  
QY 964 SSCRLYISGPNKDKFCPSTVNSL-----NCLKNCM-----SITMGSV--RPSVEOPH 1009  
Db 974 SPCKVYVHDN--TFCTNRNKSALDDKACRTCMDDPYVANSYPKSSIMYHRPSIEVFY 1031  
QY 1010 KYLPWFNLDRNRIKCPKGLAAYSTSNLTSDGVLASRFMAYHKPL--KNSQDYTEALR 1067  
Db 1032 RHLRHFLDEDTNSCEVCGGRASFKDAISFTSGRIQASQFMTHFKKLSISNSSDFIKAMD 1091  
QY 1068 AARELANITADLRKVPGTDPAFVFPYPTITNVFYEQLTILPEGLFMLSCLVPTFAVS 1127  
Db 1092 TARMVSRRLERSI-----DDTAHVFAYSKIFPPEYQYSTIMPLITLTQLFITVGVFGII 1145  
QY 1128 CLILGLDLRLSGLNLLSTVMILVDTVGFPMALWDISYNVSLINLVSAGMSEVFSVHTR 1187  
Db 1146 CVTLGIDVKGACAVICQVSNYFHVAFMIFNIPVNLASATNLVMSGIIEFSVNVLK 1205  
QY 1188 SFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIIPFFRLNLI 1247  
Db 1206 GYACSLQRADRAESTVSGIPILSGPVVTVAGSTWMLSGAHLQIITVYFFKLFLITI 1265  
QY 1248 LLGLLHGLVFLPVLISYVGDPVNPALALEOKRAEAAVAAVWVASCNPHPSRVSTADNIYV 1307  
Db 1266 VSSAVHALIITILLAFGSGRHSSTETNDNEQHDACVLS--PTAESHISNVESGIL 1323  
QY 1308 N-----HSFEGSIKGAGAINFLPNNGRQF 1332

Db 1324 NRPSSLDAHLDPLKABGGIDKAI---GRDF 1353

RESULT 4

S52525

probable membrane protein YPL006w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein LPallw; hypothetical protein yp8132.07

C;Species: Saccharomyces cerevisiae

C;Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004

C;Accession: S52525; S59687

R;Badcock, K.; Churcher, C.

submitted to the EMBL Data Library, February 1995

A;Reference number: S52519

A;Accession: S52525

A;Molecule type: DNA

A;Residues: 1-1170 <BAD>

A;Cross-references: UNIPROT:Q12200; UNIPARC:UPI000006A19C; EMBL:Z48483; NID:9683777; PII

A;Experimental source: strain AB972

R;Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa

submitted to the EMBL Data Library, August 1995

A;Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.

A;Reference number: S59677

A;Accession: S59687

A;Molecule type: DNA

A;Residues: 1-1170 <HAL>

A;Cross-references: UNIPARC:UPI000006A19C; EMBL:U33335; NID:9965076; PID:9965087; MIPS:Y

C;Genetics:

A;Gene: SGD:NCRL1

A;Cross-references: SGD:S0005927; MIPS:YPL006w

A;Map position: 16L

C;Keywords: transmembrane protein

F;1-17/Domain: transmembrane #status predicted <TM1>

F;258-274/Domain: transmembrane #status predicted <TM2>

F;341-357/Domain: transmembrane #status predicted <TM3>

F;585-601/Domain: transmembrane #status predicted <TM4>

F;614-630/Domain: transmembrane #status predicted <TM5>

F;667-683/Domain: transmembrane #status predicted <TM6>

F;698-714/Domain: transmembrane #status predicted <TM7>

F;751-767/Domain: transmembrane #status predicted <TM8>

F;1004-1020/Domain: transmembrane #status predicted <TM9>

F;1027-1043/Domain: transmembrane #status predicted <TM10>

F;1051-1067/Domain: transmembrane #status predicted <TM11>

F;1103-1119/Domain: transmembrane #status predicted <TM12>

F;1137-1153/Domain: transmembrane #status predicted <TM13>

Query Match 19.2%; Score 1329; DB 2; Length 1170;

Best Local Similarity 28.8%; Pred. No. 1.1e-84;

Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;

QY 10 LKWL-----LLRLAQSPPYTHQPCYAFYDECCKNP-----ELGSLMTLSNVSCLSN 60

Db 3 VLMIIALVGQMLRVQ-----TATCAMYNGCKKSVFGNELPCVPSPRSEPPVLSD 54

QY 61 TPARKITGDHLILLOKTCPLRYTGPNTQACSKALVSLASLITKALLTRCPACSDNF 120

Db 55 ETSK-----LLVEVCGEWEKVR-YACCTKQDVVALRDNLQKAPLSSCPACLNKF 105

QY 121 VNLCHTCSNPQSLFNVTVAQLGAGQLPAVVAVBAFYQHSFAEQSYDCSRVRVPAA 180

Db 106 NNLFCFTCAADQGRFNITKV-EKSKEDKDIIVELDFWNSSWASEFYDCKNIKPSAT 164

QY 181 ATLAVGTMCVGYSGALCNAQRWLNFOQDGTGN--GLAPLDTFHLLPQGVAGSGIQPLNE 238

Db 165 N-----GYAMDILGGAKNYSQFLKFLGDAKPMLGSPFOINIKYDLANEE--KEWQFND 218

QY 239 GVARNESQGGDVATCSQDCAACAPATAPQALDSIFYLQMPG---SLVLIIILCSVF 295

Db 219 EYVACDDAQ-----YKACSDQCESCPHL-KP-LKDGVCVKGLPCLFSLVLIPTICALF 272

QY 296 AVVTTLLVGFVRVAPARSKMVDPKGTSLSDK-----LSPSTHT-----LLGQFFQGWG 345

Db 273 AFMYYLCKRKGKNGAMIVDDDDIVPESG-SLDESENVFESFNNEFPNGKLANLFTKVG 331

QY 346 TWVASWPLTILVSVIPVVALAAGLV-FTELTTPDVELWSAPNSQARSEKAFDHOFGFP 404

Db 332 QFSVENPKILLITTVFSIFVFSFIIOYATLETDPINLWVSNSEKFEKEKEYFDNFGFP 391

QY 405 FRTNQVILTAPNRSRYRDSL--LLGPKNFSGILDLDLLELLELLELLELLELLELLELLEL 462

Db 392 YRTEQIFVFNETGPPVLSYETLHWFPDENF-----ITEEL-----QSS 429

QY 463 RNISLQDICYAPLNDPNTSLYDCCINSLLQYFQNNRTLLLLPANOTLMQGTQOVWDKHF 522

Db 430 ENIGYQDLCLFRP-TEDST---CVIESFTQYFQG-----ALPNKDSWKREL 470

QY 523 LYCANAPLTFKDGITALALSCMADYGAPVFPFLAIGYKGDYSEAEALMTTSLNNYPAG 582

Db 471 QECGKFP-----VNCLPTFOQLKTNLL---FSDDDIILNAHAFVVTLLLTWH--- 514

QY 583 DPLAQAQLWEAEFLAEEMAFQRMAGMFQVTFATERSLEDEINRTTAEDLPFATSYIV 642

Db 515 --TQSANWEER-LEEYLLDLKVPBGL-RISFNTSISLEKELNN--NDISTVAISYLM 567

QY 643 IFLYISLALGYSYSGRWVVDKATLGLGCVAVVVGAVMAAMGFFSYLGRSSIVLQV 702

Db 568 MELYATWALRRKDKTLL-----LGISLLIVLASIVCAAGFLTLFLGLKSTLIIEAVI 621

QY 703 PFLVLSVGADNIFIFVLEYQR-LPRRGPGEPRVHIGRALGRVAPSMLLCSLSAICFLG 761

Db 622 PFLILAIGDNIPLITHEYDRNCEQPEYSIDQKIISAIGRMSFSLMSLLCQTGCFLIA 681

QY 762 ALTPMPAVRTFALTGLAVILDFLQMSAFVALLSLDSKQEAASRLDVCCKVPQELPPP 821

Db 682 AFVTMPAVHNFAYISTVSIFNGVLQTAIVSILSLYEKRSNYKQIT--- 728

QY 822 QGSEGLLGFQKAYAPFLHWTIRGVVLLFLFALFGVLSYSCMCHISVGLDQELALPKDS 881

Db 729 -GNEETKESFLKTYFKMLIQ---KRLIIIFSAWFTSLVFLPEIQGLDQTLAVPODS 784

QY 882 YLLDYFLFLNRYFVGAPVPTVTLGVNFSSEAGNNAICSS-AGCNNPSTFKQIQADEF 940

Db 785 YLVDFYKDVYSFLNVGPPVYVVK-NLDLTKRQNOQKICGKFTTCERDSLNVLE---QE 840

QY 941 PEQSYLAIPASSWDDFDIMLTPSS--CCRLYISGPNKDKFCPTVNSLNCNKMSITM 998

Db 841 RHRSTITEPLANWLDYIPMLNPQNDQCRL---KKGIDDEVCPSPFRRRC-ETCFQ--Q 894

QY 999 GSVR-----PSVEQFHKYLFWLMDRPNICKPKGLAAYSTVNLSLDGQVLAERFAY 1052

Db 895 GSWNTNMSGFPPEGKDFMEYLSIWIN-APSDPCPLGGRAPYSTAL-VYNETSVSASFRTA 952

QY 1053 HKPLKNSQDYTBALRAARELANIYADLRKVPGTDPAPFVRPYTITNVFYEQVLTILPBG 1112

Db 953 HHPLRSQKDFIQAY-----SDGVRISSSSPFLDMFAYSFYIFVQYQTLGLPT 1001

QY 1113 LFMLSLCLVPTFAVSCLLGLDLRSGLNLLSLVMILVDVTGFMALWDISYNVSLINLV 1172

Db 1002 LKLGSAIILIFISSVFL-QNIRSFLLALVVTMIIVDIGALMALGSLNAVSLVNI 1060

QY 1173 SAVGMSVEFVSHITRSFAY---STKPTWLERAKEATISMGSAVFAGVATNLPGLVLGL 1229

Db 1061 ICVGLGVERCVHIVRSFTVPVSETKDKANSRVLYSLNTIGESVIKGITLTFKFGVCVLA 1200

QY 1230 AKQAQLQIFPFRNLNLTLLGLLHGVLPVLSYVG 1266

Db 1121 AQSKIFDVYFRMWFLLIIVAAHLHLLFLPALLSLFG 1157

RESULT 5

T30172

transmembrane protein patched - mouse

C;Species: Mus musculus (house mouse)

C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T30172

R;Goodrich, L.V.; Johnson, R.L.; Milenkovic, L.; McMahon, J.A.; Scott, M.P.



Mon Apr 10 07:16:16 2006

388 NEDKAAAILLEAQRMTVEVYVHOSVAQNSTQKVLSTTT--TTLLDDILKSPSDSVIRVASGY 446  
 Db  
 641 IVIFLYISIALGYSYSSWVMDSKATLGLGVAVVGLGAVMAAGFFSYLGRSSLLVILQ 700  
 QY  
 447 LLMLAYACITM---LRWD--CAKSOAGVLAGVLLVALSVAAGLGLCSLIGISFNAATQ 501  
 Db  
 701 VVPEFLVSGADNIFIVLEVORLPRRPGEPREHIGRALGRVAVPSMLLCSLSEALCEPL 760  
 QY  
 502 VLPFLALGVGVDDVFLLAHAFSETGQNKRIPEEDRTGECIKRTKGASVALTSISNVTAFFM 561  
 Db  
 761 GALTMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLDSKROEASRLDVCCC----- 812  
 QY  
 562 AALIPALRAFSLQAQVVVVFVAVMLLFFPAILSMDLVRRDRDLDFCCFTSPCVTR 621  
 Db  
 813 ---VKPQ---ELPPGGGEGCL----- 827  
 QY  
 622 VIQIEPOAYAEADNICVSSPPSYSSHFAHETQITMSTVOLRTEYDPHTQAYYTTAEPR 681  
 Db  
 828 ---LLGFQKAYAP 838  
 QY  
 682 SEISQVPTVTDLSLSCSPESASSTRDLSLQSFSDSVHCLPEPCTKWTLLSTFAEKHYAP 741  
 Db  
 839 FLLHWITRGVLLFLALFGLVSLYSMCHI SVGLDOELALPKDSYLLDYFLFLNRYFEVGA 898  
 QY  
 742 FLLKPKAKVWVIFLFLGLLGLSLYGTTRVRDGLDLDIVPRDITREYDFIAAQKYFSF-- 799  
 Db  
 899 PVYFTTLYGNFSEAGNAICSSAGCNFFSFTQIOYAT-----EFPEQSILAI 948  
 QY  
 800 ---YNN-----YIVTQADYPNVQHLLYELHRSFNVTVYLL 833  
 Db  
 949 PA-----SSWDDDFIDWL-----TPSS-----CCLYISGNKD 977  
 QY  
 834 EGRDLQPKWMLHYFRDMLQGLQAFDSWETGKITYSNYKNGSDDAVLAYLLVOTGNRA 893  
 Db  
 978 KFCPTSVNSLCLNKMSTMGSVRPSVEQPHKYL- PWFLND-----RPNI 1022  
 QY  
 894 K-PIDISQLT--KORLVDADGIINPNA--FYIYLTAWNSNDPVAASAQANIRPHRPEW 947  
 Db  
 1023 KCPKGGLAAY--STSVNLSDGOVLASREMAVHKPLKNSODYTEALRAARELAANITA-D 1079  
 QY  
 948 VHDK---ADYMPETRURIFAPRIEYAOFPFYINGLRETSDFEAIEKVRACNNYTSIG 1004  
 Db  
 1080 LRKVPCTDPAFEVFPYTIITNVFEYQYLITLPEGLFMLSCLIVPTFAVSCLLGLDLRSLG 1139  
 QY  
 1005 IASYPNG-----YFP-----LFWEQVIGLRHMLLSISVVLACTFLVLCALFLNPTAGI 1054  
 Db  
 1140 LNLISVIMILVDTGVFMALWDISYNAVSLINLSAVGMSVEFVSHITRSP--AISTKPTW 1197  
 QY  
 1055 I-VVVLALMTVELFGMGLIGIKLSAVPVVILLIASVGIGVEFTVHALAFLTAIGDKN-- 1111  
 Db  
 1198 LERAKEATISMGSAVPAVAGVAMTNLPGILVLGLAKAQIQTFFRLNLITLLGLHLGVF 1257  
 QY  
 1112 -RRAVTALAHMFAPVLDG-AVSTLLGVLMAGSEDFIVRYFFAVLAILITLGLVNLGL 1169  
 Db  
 1258 LPVILSVYG--PDVNPALA---LEOKRAEAAVAAVASCNPHSPRVSTADNIVNHSPE 1312  
 QY  
 1170 LPVILSFGPYPEVSPACGRNRLTPSPPEPPPSIVRFALPPGHTNGS--DSSDSYSSQ 1227  
 Db  
 1313 GSIKG 1317  
 QY  
 1228 TVVSG 1232  
 Db

RESULT 7  
 S44797  
 F09G8.4 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 14-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
 C:Accession: S44797  
 R:Anderson, K.  
 submitted to the EMBL Data Library, February 1993  
 A:Description: Sequence of the C. elegans cosmid F09G8.  
 A:Reference number: S44776

A:Accession: S44797  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-915 <AND>  
 A:Cross-references: UNIPROT:P34389; UNIPARC:UPI000013B859; EMBL:L11247; NID:G156280; PII  
 C:Genetics:  
 A:Introns: 51/1; 214/3; 255/3; 326/3; 382/2; 409/3; 455/3; 505/3; 569/2; 702/3; 836/3

Query Match 8.7%; Score 600; DB 2; Length 915;  
 Best Local Similarity 22.7%; Pred. No. 8.9e-34;  
 Matches 236; Conservative 158; Mismatches 422; Indels 224; Gaps 37;

QY 10 LLWALLLRLAQSEPTTHIQPCYCAFYDECNGNPNLSGLMTLSNVCSLNTNPARKITGD 69  
 Db 14 VLFLLLIHLAQ-----AKVM--TECDGEEDSNHPPCKTNKSTVLPITVTRSLNPT 64

QY 70 HLILQKICPLRYTGN--TQACCSAKQLVSLBASITKALLTRCPACSDNFVNLCHNT 128  
 Db 65 YNARFEKYSYLVQEDKAQVCCBELQLKGMDRISNAATILGSCPCDFNFAXLWCQFT 124

QY 129 CSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYCHSFAEQSYDSCSRVRVPAATAATLAVGTM 188  
 Db 125 CSPDQSKFMKVMET-----TGPKNVVVMKMEFKVNRDFEGLYESCRHTWANGLAFLMSL 180

QY 189 CGVGSALCWAQWLNFGDGTG--NGLA---PLDITFHLLEPGQAVGSGIOPLNEGVARCN 244  
 Db 181 GGRKVS-----FENFYGMGTKNLAQSIPIINTEFQFSRMKNAMNIPPTP-----CH 225

QY 245 ESQDDVATGSCQDC---AASCPAIPARQALDSTFFYLGQMPGSLVLIILILCSVPVVUTIL 301  
 Db 226 KSAQPKVPACAGIDCPTNAHQVLVDISKVEHLGTVKVFHPFP--DFEWLLKICGCLA-LTVL 283

QY 302 LVGF-----RVAPARDKSKMVDPKGTSLSKLSFSTHTLLGQFPQ-----WGTWVAS 350  
 Db 284 LVFTLKYSCHRRSAPNGEDGCVYDLGKGN-----LEVQEGLCARVANAVIK 330

QY 351 WPLTILVLSVTPVVALAAG--LVFTLTPDPVWLSAPNSQARSEKAF-HHQHQPFPFRTN 408  
 Db 331 HPLIFVSLGLLIVAAACCGNFKFHSLSHTSDVQSAADGETRNEKKETIHS--FGPNRIE 388

QY 409 QVILTAPNRSSRYDLSLLGLPKNFSGILDLDLLELLEQLERHLQVWSPQARNISLQ 468  
 Db 389 QIFNLPTT-----KSMFNMPLFEMFQLVGNIQNLT--ACYGNSSVKLD 432

QY 469 DICYAPLNDNTSLVDCINSLLOFQFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANA 528  
 Db 433 DICYKPIGKN-----HGCIMSPNTNYFO----- 455

QY 529 PLTFKDGITALSCMADYGA PVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQ 588  
 Db 456 -----YYTARTIMITVLIING-----PE-DQ 474

QY 589 AKLWBEAFLEEMRAFORRMAGMFOVTTAERSLEDEINRTAED--LPFATSYIVIFLY 646  
 Db 475 ALAWETAFLNMSRYEMKHANF---TFMTETSAEEIHTAVETDKITVSVIACNAVLWVI 531

QY 647 ISIALGSY--SSWSRVWVDSKATLGLGGVAVVGLGAVMAAGFPYSLGRSSLLVILQVWP 704  
 Db 532 TMLGINHPRESSILSALVHHKLLISISAVMISVISVWCISGMSLFGVHATDNAVIVLFF 591

QY 705 LVLSVGADNIFIFVLEVQRLPRRPGEP-----REVH--IGRALGRVAPSMLLCSLSEALCF 758  
 Db 592 VITCLGINRIFVIRTFIQANGHCYGLPNISYREMNHRISNVMRSPVILVTLNLSICSTCL 651

QY 759 FL-GALTP-----MPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKRO----- 802  
 Db 652 FLAGGVLPYVSVSMPAVEVEARHAGLAIMLMDTAFYLLVMLPLFOYDARREMSGKEIWPW 711

QY 803 -----EASRLDVCCCKVQDELPPQGGEGLLLPFPQKAYAPFLLHWITRGVLLFLALFG 858  
 Db 712 YELSNESKINLCMEAVDGNLRSP-----VDWFKLAIAPLLLAKKICRIWIATFFVSLI 764

QY 859 VSLYSMCHISVGLDQELALP-----KDSYLLDYFLFLNRYFEVGA PVYFT 904

```
Db 765 IACVCTLCLEFGFNQWAFSEVCIFPIFSLHNGSFQTSYLTTHFQNMENLNIIGPLPFWV 824
Qy 905 TLGYNFSSEAGMAICSSAGCNFSFTQKIQ---YATEFPPEQSYLAIPASSVWDDFDLWL 961
Db 825 EGDVKWHPDKMKNKFTLAGCDDNSMGKIRSLAYAEY-KGNVYLHGDVNIWLDLSYLQPM 883
Qy 962 TP-SSCCRLYISGPNKPKFC 980
Db 884 HPRGSCCKM-----DGKQFC 898

RESULT 8
T25600
hypothetical protein C32E8.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25600
R;Gattung, S.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid C32E8.
A;Reference number: Z20056
A;Accession: T25600
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-933 <GAT>
A;Cross-references: UNIPROT:P91129; UNIPARC:UPI00000611CF; EMBL:U88308; PIDN:AAB42325.1.4
A;Experimental source: strain Bristol N2; clone C32E8
C;Genetics:
A;Gene: CESP:C32E8.8
A;Map position: 1
A;Introns: 145/1; 177/2; 385/1; 418/3; 833/3; 878/3

Query Match 8.74; Score 598.5; DB 2; Length 933;
Best Local Similarity 22.98; Pred. No. 1.2e-33;
Matches 251; Conservative 181; Mismatches 362; Indels 303; Gaps 37;

Qy 341 FQGWGTWASWPIITLVSVIPVVALAAGLVFTLTTPVELWSAPNSQARSEK-AFH-- 397
Db 14 FROLGLICDHPPLFPFFVPLFTAAAGVGLHLNPLSDAVYLTPLGAQSKMERMSIHEK 73
Qy 398 ----DQHFGP---PFRTNQVILTAPNRSRYSDLSLGPKNFGSLDLDLLELLEQER 450
Db 74 WPUTDNNYIPGRAVTSREIQVFTALARN-----DSNILDPKFANAVYQLDKY-----IQTR 124
Qy 451 LRHLQVMSPEAQRNISLQDICYAPLPNDNTSLYDCCINSLLQYFQNNRTLLLTANQTLM 510
Db 125 VRVLH----- 129
Qy 511 GOTSQVDWKDHFPLYCANAPLTFKDGTALE-----LSCWADYG-----APVFPFLAIGGY- 559
Db 130 -----NGHYYSYKNLCLOYKNGGCPSNKHVHILSDLNHHGNITYPVYFRFGSEGGYI 181
Qy 560 -----KGXDYE-----AALMTWLSLNYP-----ACDPLAQAKLWEAFLEE 599
Db 182 GSSLGGVTVMKGENETDILASAKAWPMIYHLKHPBEMSYSISG-----WE-----LEL 230
Qy 600 MRAFORRMAGMF-QVTFTRASRSELEINRTTAEDLPIEATSYIVIFLISLALGSY--SS 656
Db 231 GRMLTQVPEDPYISITYFISQTLADELKNWADTLIPFIISITLLIVFSLCSLSFDGS 290
Qy 657 WSRVMVDSKATLGGGVAVVLGAVMAAMGFYSYIGIRSSLIVLQVVPFVLVSGADNIFI 716
Db 291 FSIDVWLKPFILSILGVWSAGAILTGVGFLSLMGMPYN-DIVGVMPFLVLAAGVDNMFL 349
Qy 717 FVLEYORLPRRCPPEVHVH--IGRALGRVAPSMMLCSLSAICFFFLGALTPMPAVRTFAL 774
Db 350 MVAAV-----RRTSRTHVHERMGECLADAASVILITSSTDVLSFGVGAITTIPAVQICV 405
Qy 775 TSLGAVILDFLLQMSAFVALLSDSKRQEAR-----LDVCCCVK----- 814
Db 406 YTGVAITFAFFIYQITFFAACLALAMKHEASGRNSLFLIEAVSAEKTSISTFORLNLGS 465
```

## RESULT 9

T13952

membrane protein ptch2 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T13952

R;Motoyama, J.; Takabatake, T.; Takeshima, K.; Hui, C.

Nature Genet. 18, 104-106, 1998

A;Title: Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog.

A;Reference number: Z17830; MUID:98122566; PMID:9462734

A;Accession: T13952

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1182 &lt;MOT&gt;

A;Cross-references: UNIPROT:Q35595; UNIPARC:UPI0000001587; EMBL:AB010833; PIDN:BAA2469

A;Experimental source: strain BALB/cCrSlc

C;Genetics:

A;Gene: ptch2

C;Superfamily: Drosophila membrane protein patched

C;Keywords: transmembrane protein

Query Match 8.44; Score 583.5; DB 2; Length 1182;

Best Local Similarity 24.24; Pred. No. 1.8e-32;

Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

Qy 330 SFSTHTLLGQ-----FFQG-----WGTWVASWPLTILVLSVIPVVALAAGLVFTLT 376

Db 20 SSAPHILAGSLQAPLWLRAYFQGLLFLSLGCRIOKHGKVLFLGLVAFGALGLRVAVIE 79

Qy 377 TDPVELWSAPNSQARSEKAFHQHFG-PFFRTNQVILTAPNRSRYSDLSLGPKNFSGI 435

Db 80 TDLEQLWVEGSRVSGELHYTKELGEEAAYTSQMLI-----QTAHQEGGNVLTPE----A 131



```
QY 996 ITWGSVRPSVEQFHKYLPHFLNDRPNKCPKGGGLAAYSTSVNLTSQGVLT-ASRFMAYHK 1054
Db 630 -----VFLSQPQF-----AKYNRDVVLTEDEGLEASRI----- 658
QY 1055 PLK-----NSQDYTEALRAARELANATADLRKVPGTDPAFEPVPPYITNVFQYLTILP 1110
Db 659 PVQLRHVGANSRWRLLFRRLAE--TSELQ-----TGVYADFFQ-----FAEQYNAVLP 706
QY 1111 EGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIVMILVDTVGFWMALWDISYNAVSLIN 1170
Db 707 GTLSSIAVAGVAVAVSLILPEPVAS-LWVSFVSINIGILGFMTFWSVRLDFISWT 765
QY 1171 LYSAVGMSVEFVSHITRSFAITSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIIVGLA 1230
Db 766 IVMSIGFCVDFAAHLAYNFAGKQNMDSRMRNALYAVGAPILMS-ATSTIIGVSFMSA 824
QY 1231 KAOLIQIPFRNLTLITLGLLHGLVFLPVILS--VVGPDVNPALALQKRAEEA- 1287
Db 825 ESYVFR-SFLKTMIVLILGALHGLVILPVLLSMFYCGG-----SSKKAKEHIDAVD 875
QY 1288 -MVASCPNHPGRVST 1301
Db 876 QKLQAYNNPARTAS 890

RESULT 11
Ti8291
patched protein - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: Ti8291
R:Concordet, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.;
Development 122, 2835-2846, 1996
A:Title: Spatial regulation of a zebrafish patched homologue reflects the roles of sonic
A:Reference number: Z18860; MUID:96379744; PMID:8787757
A:Accession: Ti8291
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1220 <CON>
A:Cross-references: UNIPROT:Q98864; UNIPARC:UPI0000132883; EMBL:X98883; PIDN:CAA67386.1
C:Genetics:
A:Gene: Ptc1
C:Superfamily: Drosophila membrane protein patched

Query Match 8.3%; Score 572; DB 2; Length 1220;
Best Local Similarity 22.4%; Pred. No. 1.2e-31;
Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;

QY 321 KGTSLSDKLSFSTHTLLGQPFQGGTGWASWPLTILVLSVIVVVALAAGLVFTBLTDPV 380
Db 51 KGKAVGQKAPLWIRARFOAPLFLSLGCHIQRHCGKVLFIGLLVFGALSGLRVAIETDIE 110
QY 381 ELASAPNSQARSEKAFHDQHFQ-PFRFNQVILTAPNRSSVRYDSLLGKPNFSGIILDD 439
Db 111 KLMVEAGSRVSKELRYTKERQGESVFTSQMLIQTPKQEG-----TNILTOE 157
QY 440 LLELELEQLRHLRHLQVWSPQAQRNLSLODICVAPLNP--DNT-----SLYDCCINSL 490
Db 158 AL--LLHLEALASKQVSVLSYKGSWDLNKLKCFKSGVPIIENVMIERMIDKLPFCMIIVTP 215
QY 491 LQVFQNNRTL-----LLLTANQ--TLMGQTSQVDWDKDFLYCA 526
Db 216 LDCFWEGSKLQGSAYLPGMPDIOQMNLDPLKMEELSQFTSLGPFREMLDKAQVGHAYM 275
QY 527 NAPLTFKDGKTALALSCHADYQAP-----VFPFLAI-----GGYKG----- 561
Db 276 NRPECLDPSDT-----DC--PHSAPNKPQWQVNTAAELQGGCHGFSKFKFMHQBELILGER 329
QY 562 -KD-----YSAEALIMTFSL-----NNYPAGDPRLAQAKL-----WEEAPLEE 599
Db 330 VKDSQNALQSAEALQTNWFLMSKPLQYEHFKDDYIEHDINWEDKATAILESQWRKFVEV 389
QY 600 MRAP--ORRMAGMFQVTFPTAERSLEDEINRTAEDLPFATSVIVIFLYISLALGSYSSW 657
```

## RESULT 12

T29590

hypothetical protein F55F8.1 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29590

R:Gattung, S.; Scheet, P.; Kemp, K.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of *C. elegans* cosmid F55F8.

A:Reference number: Z20647

A:Accession: T29590

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

```
Db 390 VHGSIPQSSNNVAFSTI---TLNDIMKSFSDSVIRVAGGYLLMLACTVM---LRW 443
QY 658 SRVMVDSKATGLGCVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIF 717
Db 444 D--CAKSOQAGVAGVLLVALSVAAGLGLCSLLGLSFAAATTVQLVPSLAGIGVDDMFL- 500
QY 718 VLEYQRLPRRCPGEPREVHIGRALGRVAPSMILCSLSAICPFLGALTMPAVRTTALTSG 777
Db 501 -LGHSTTETRNIIPKERTGDCRLRTGTSTVALTSVNNMIAFFMAALVPIPALRAFSLQAA 559
QY 778 LAVILDFLLMSAFVALLSLDSKROEASRLDVCCC-----VKPQE----- 817
Db 560 VVVVFNFAWALLIPALLSLDLHREDRKLDILCCFYSCSRVLIQIQOELSDANDNHQ 619
QY 818 -----LPPGQ----- 823
Db 620 RAPATPTYGTSTITTTTHITTTTVOAFTOCDAAQGHIVTILPPTSQISTTPPSMVLSTPTP 679
QY 824 -----GEGE-----LLGFFQKAYAPFLHLHWTR 846
Db 680 TTDYGSQVFTTSSSTRDLLAQVEEPKEGECVLPFFFRWNLSSFAREKYAPLLKPKETK 739
QY 847 GVLLLLFLALFGVSLYSMCHISVGLDQELALPKDSYLDYFLFLNRYEVEVGAPVVFVTTL 906
Db 740 TVVVVVALLSLSLYGTTMVHDGLYLDIVPRDQOEYEFITAQKYSF-YNNMLVLTWD 798
QY 907 GYNFSSEAGMNAICSSAGCNNFSPQKIQYATEPEQSYLA-----IPASSWVDDFDIW 960
Db 799 GFDYAR-----SQQLQLHNAFNVSVKYVVKDGNHKLKLP-RMWLHYFQDW 841
QY 961 LTPSSCC-----RLYISGPNKDKFCBPSTVNSLNLKNC 993
Db 842 LKGLQATFDADWEAGKITYDSYRNGTEGALAYKPLIQTSKGKPEFNYSQLTSSRLVDG- 900
QY 994 MSITWGSVRPSVEQFHKYLKLP-WFLNDRPNKCPKGGGLAA-----YST--- 1034
Db 901 ----DGLIPEV--FYIYLTVWSND-----PLGYAASQANFYPHPREWHDKYDTTGE 948
QY 1035 SVNLTSQGVLASRFMAHYHKLKNSQDYTEALRAARELAN-ITADLRKVPGTDPAPFVF 1093
Db 949 NLRIPAAEPLFAQFPFLNGLRQASDFIEAIESVRTICEBFMRQGIKNYENG-----Y 1002
QY 1094 PYTITNVFQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIVMILVDTV 1153
Db 1003 PF-----LFWQYIGLRHWHFLLSISVWLACTFLVCAILLLNPWTAGVI-VFILPMWTVELF 1057
QY 1154 GFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSF--AISTKPTMLERAKEATISGSA 1211
Db 1058 GIMGLIGIKLSAIPVVLIIASVGLGVEFTVHIALGFLTAIGDRNT---RSVAMEHMFAP 1114
QY 1212 VFAGVAMTNLPGLIIVLGLAKAQIIFIFFRLNLLITLLGLLHGLVFLPVLSYVGP--DV 1269
Db 1115 VIDG-AISTLLGVMLAGSEDFIMRYFFAVLAILTLGLINGLVLPLVLLSLMGPPAEV 1173
QY 1270 NPA 1272
Db 1174 VPA 1176
```

## RESULT 12

T29590

hypothetical protein F55F8.1 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29590

R:Gattung, S.; Scheet, P.; Kemp, K.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of *C. elegans* cosmid F55F8.

A:Reference number: Z20647

A:Accession: T29590

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

Db 783 ARVCDALGDLGWPVAQG-ANSTILAVSLSDVPAYMI-VTFKTVFLAISIGFLHGLVFL 840  
QY 1259 PVILS-YVG 1266  
Db 841 PLMLSVFVG 849

RESULT 13

T27969  
hypothetical protein ZK675.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27969  
R:Stms, M.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: Z20448  
A:Accession: T27969  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1405 <WIL>  
A:Cross-references: UNIPROT:Q09614; UNIPARC:UPI00000835A5; EMBL:Z46812; PIDN:CAA86843.1  
A:Experimental source: clone ZK675  
C:Genetics:  
A:Gene: CESP:ZK675.1  
A:Map position: 2  
A:Introns: 31/1; 59/3; 100/3; 244/2; 485/1; 950/2; 1298/2; 1335/3  
C:Superfamily: Drosophila membrane protein patched

Query Match 7.4%; Score 514; DB 2; Length 1405;  
Best Local Similarity 20.2%; Pred. No. 1.7e-27;  
Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLDGLSFSHTLLGQFFQCGTWV--ASWPLTILVLSVIPVVAALAGLVFTLTDD 378  
Db 100 RGAATGNRYALYSRSLIKLLFALGNTVHRNAWSI-ILAVSMIFAVC-CVGLQYVHIETD 157  
QY 379 PVELWSA-----PNSQARSEKAFDQHGFFP-RTN-----QVILTAENRSS 419  
Db 158 IVKLVAQAQGRLEENFLNFKAMRNVTGDS--GPPLPRENGLGGGVQVLTOTPEYEG 215  
QY 420 YRVDLLGLPKNFSGIILDLLELLELQERLHLQVMSPEAQNI-----SLQDICY 472  
Db 216 --QDLAAGP-----LLKHEVIMKHIAFNVSVHGVDSLSIDCF 253  
QY 473 APLNPD-----NTSLVD-----CC-INSLLQYFQNNRTL-----LLL 503  
Db 254 KPAPPSVAADSAASSLGVIDKIVPCIWITPDCFWEGSKALGPHPSLPKSSSLGFLGMLL 313  
QY 504 TANQTLMGTSQVDWKD-----HFLYCANAPLTP----- 532  
Db 314 SS-----LSDGDMIRNSDFDPIAVIDEIHRSFNLSGSHYTFERAGVSHGYMDRPCIDPLDP 369  
QY 533 -----KDGITALALSCWADYGAPVFPFLAIGYKGDYSRAE 568  
Db 370 ECPPMKKNYFDVCPHIDRVREIAKYGTLEBEKKKSGYSFDFL--GRKKREAGDOP 426  
QY 569 ALI----- 571  
Db 427 KMIHPAQPADSIPTIEDAVPAQVPVSTAPIPTTTLSPEARAAAEKKEKKQKARELKDYC 486  
QY 572 -----MTFSLNYP-----AGDP 584  
Db 487 KSYKSAFELKKKDKKPEVMSSENYPNQVNDYAAEWTCGCGFASNLVNPEDMILGNP 546  
QY 585 RLAC--AKL-----WEEAFLEE-MRA 602  
Db 547 RRKKGGKLGADALQSVFLVSPADVFLRFKQKPGNSMKTGLDMDANWETAABQVLOA 606  
QY 603 FORRMAGMFQVTFTEARSL-----EDEINRTTAEDLPIFATS-----YIV 642  
Db 607 WQR-----NFT--KSLYNHKNVDEGNGRERLT--PLASTSIADMLEEFQCFNYTI 654

A:Residues: 1-889 <GAT>  
A:Cross-references: UNIPROT:P91346; UNIPARC:UPI0000178A20; EMBL:U80447; PIDN:AAB37812.1;  
A:Experimental source: strain Bristol N2; clone F55F8  
C:Genetics:  
A:Map position: 1  
A:Introns: 36/3; 66/3; 98/3; 149/3; 200/3; 240/2; 279/1; 358/1; 481/2; 569/3; 606/3; 697/3

Query Match 7.8%; Score 539.5; DB 2; Length 889;  
Best Local Similarity 23.0%; Pred. No. 1.5e-29;  
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;

QY 401 FGPFPPTNOVIL-----TAPNRSYRYSLLGPKNFSGLIDLLLELLE 446  
Db 31 FGYSTYERRIHDAMPVLDGTFVAGRAVTS--REVQAVVARGSGNILDVRFVSNELKL 88  
QY 447 LQERLRLHQSPEAQNISLQDICYAPLNPNNTSLYDCCINSLLQYFQNNRTLLLTAN 506  
Db 89 MESFIRN-NITVQFSNRTWSFADCLLA--GPDGR-----CANNDHIQ----- 127  
QY 507 QTLMGQTSQVDKDHFLYCANAPLTFKDGTPALALSCWADYGAPVFPFLAIGGVK----- 560  
Db 128 --LASRLHQ-----HGINITYPTVRLSDKSAIYAS-----ALGGVKLAKGN 167  
QY 561 GKD-YSEAEALIMTFSLNYP-----AGDPLRAQAQKLWEEAFLEENRAFQRMAGMFOV 613  
Db 168 GENIIVEATAMLLIYQLKFENEISYVSG-----LWREFPKWDEY-KKQAKYISI 218  
QY 614 TFTAERSLEDEINRTAEDLPIFATSIVIFLY-----ISLAGSYSSWSRVWDVKAT 667  
Db 219 TYFHSQTLSDLENNRANERLAKFKFAGFVILCFVSLGSIIVTKSGYIDW---VTKPI 274  
QY 668 LGLGVAVVLGAVMAAGFFSYGLIRSSILQVFPFLVLSVGADNIFIFVLEYQRLPFR 727  
Db 275 LSVLGVSNAGMGIASAMGMLTYLEIQYN--DIIANVPFLVAVGTDNMFVWASLKRTRD- 332  
QY 728 PGEPRVHTRGALGRVAPSMILCSLSEACFPFLGALTMPMPAVRTFALTSGLAVILDFLQ 787  
Db 333 -NLKYQRIAEACWADAASVILITALTDLFSFGVGTITTPAVQVIFCIYTMCALLTFAYQ 391  
QY 788 MSAPFALLSLDSKRQE-----ASRLDV---CCCVKPOQLPPPGOG----- 824  
Db 392 LTFECALLVYTRIEBGLSHIWLRPVAVTYSSTPLNKLFWLGSQPKR-PLPSCGTVSS 450  
QY 825 -----EGLLGFFQKAYAPFLH-WITRGVLLFLALFGVSLYS 863  
Db 451 TSSVSTWTSQATSPASKHLHCAATSPFRWYAPVLMPQWI-RAIAGLWYLIYLGISIYG 509  
QY 864 MCHISVGLDQELAPKDSYLLDYFLNLR-YFEVGAPVYFVT-----LGYNFSSE 913  
Db 510 CTHLKEGLEPANLLVDSDYATPHYRVLEKHYWHYGASLIQIVSNPPDLRDPVERINMDKM 569  
QY 914 AGMAAICSSA-GCNNFSF-----TQKIQYATE-----FPEQSYLAIPASSWDDFI 958  
Db 570 ASTFANCKVAIGDSDSVQFWLREMVMQVSEIHKIQYDNEKPYDHAAQYIYSDMSQVWVDV 629  
QY 959 DMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMCSITMGSVRPSVSEQFHLYLPWFIND 1018  
Db 630 -----WGRNN 634  
QY 1019 RPNIKCPKGLAAAYSTVNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELANITA 1078  
Db 635 -----NSERIILKTPFRMIGMEDISTTKTQTEANTTFREIASRF-- 672  
QY 1079 DLKRVFGTDPAFVFPYTTNVFYEQYLTILBGLFMLSCLVPTFAVSCLLGLDLRSG 1138  
Db 673 -----RQYNVTYTMPLWLTQDQYALVVPVNTMQDIIIVAVACMLVISALLIPQVCSF 723  
QY 1139 LNLLSIVMLVDTVGPMALWDI:SYNAVSLINILVSAVMSVEFVSHITRSPALSTKPTWL 1198  
Db 724 WV-AVTIGSIDLGLVGFNTLNNVNLDAISMITIIMSVGFSVDYSAHITVAYVISKSTTS 782  
QY 1199 ERAKEATISMGSVAVFAGVAMTNLPGLIIVLGLAKAQLIQIFFRRLNLLITLLGLHGLVFL 1258



Db	857	MVTIDIGVIGLSLWSKLPDPISMITIISIGFSIESAHTGFSVNSLNSAFRCVD	916	QY	522	FLYCANAPLTFKDGKTALALSCM	---	ADYGAPVFPFLAIG--GKY	560
QY	1204	ATISMGSAVAGVAMTWLPGILLVLGLAKAQLIQIFFRFLMLLTLLGLLHGLVFLPVILS	1263	Db	265	EQMKRAAIA--SGYMEKPCLNPLNPNCPDPAPNKNSTQPPDVGA	---	ILSGGCGYGA	317
Db	917	AMEKLAWPVVHG--SLSTILGVFLAFIDSVMVLVFFKTIISLVL-IIGAMHALMLLPILLS	974	QY	561	GKDYSEAEALIMFTSLNNYPAGDPLAQA	---	---	589
QY	1264	YVGP-----DVPNPALALEQKRAEVAAVVAVASCP	1293	Db	318	AXHMHPELIVGGAKRN-RSGHLRKAQALQSVQQLMTEKEMYDQWQDNKYVHHGLWTOE	---	---	376
Db	975	MCIFVIERLSDAKSKASDRRKLKNKNVVAINLP	1010	QY	590	-----KLWEAP--LEEMRAFORRMAGMFOVTTFAERSLEDEINRTTAEDLIPATSY	---	---	640
				Db	377	KAAEVLNAWQRFNSREVEQLLRKQRIATNYDIYVFSSAALDDILAKFSPALSIVIGV	---	---	436
RESULT 15				QY	641	IVIFLYISLALGYSWSRVMDSKATLGLGGVAVVLGAVMAAMGFYSYSGIRSSVLILQ	---	---	700
S06119				Db	437	NTVLY--AFCTLLRW-RDPVGGSSVGVAGVLLMCFSTAAGLGLSALLGIVFVNAASTQ	---	---	492
membrane protein patched - fruit fly (Drosophila melanogaster)				QY	701	VDFVLVSVGANIFIVFVLEYQRLPRRPOCEPREVHIGRALGRVAPSMLLCSSEALCFEL	---	---	760
C:Species: Drosophila melanogaster				Db	493	VVPFLALGLVDHIFMLTAAYAESNR--EQTKL-----ILKKVGPISILFSACSTAGSFFA	---	---	546
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004				QY	761	GALTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKROBASRLDV--CCCV-----K	---	---	814
C:Accession: S06119; A33468				Db	547	AAFIPVPAUKVPCLQAATVWCSNLAALIVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ	---	---	606
C:Nakano, Y.; Guerrero, I.; Hidalgo, A.; Taylor, A.; Whittle, J.R.S.; Ingham, P.W.				QY	815	POELPP-----PQGEGL-----LLGPF	---	---	832
C:Nature 341, 508-513, 1989				Db	607	PKVAPPVPLNANNRGARHPKSCNNRVVALPAQNPLLEQRAADIPGSSHSLSATFA	---	---	666
A:Title: A protein with several possible membrane-spanning domains encoded by the Drosoph				QY	833	QKAYAPFLHWTGRVVVLLFLALFGVLSYMCCHISVGLDDELALPKDSYLLDYFLFLNR	---	---	892
A:Reference number: S06119; MUID:90015164; PMID:2797178				Db	667	FQHYTPFLMSRWKFLTVMGFLAALISLSTYSTRLOQGLDIIOLVFKDSEHFKFLDQTR	---	---	726
A:Accession: S06119				QY	893	YEEVGAPVYFVTTLGYNFSSEAGNNAICSGAGCNFSFTQIQIYATEFFPQSYLAIP---	---	---	949
A:Status: not compared with conceptual translation				Db	727	LF--GFSYMTAVTQG-----NFEYPTQQLLRDY-HDSFVRVPHVI	---	---	764
A:Molecule type: DNA				QY	950	-----ASSWVDDFDLWTFSSCCRLYISGPNKDFPCPSTVNSLCKNC-----	---	---	993
A:Residues: 1-1299 <NAK>				Db	765	KNDNGGLPDPWLLLFSEWL-----GNLQKIFDEEYRDGRLTKCWFNASSDA	---	---	812
A:Cross-references: UNIPROT:P18502; UNIPARC:UPI000014EA18; GB:X17558; MID:g8389; PIDN:CA				QY	994	-----MSITMGSVRPSVEQ-----FKYL--PWFLNDRPNIKCPKG	---	---	1028
R:Hooper, J.E.; Scott, M.P.				Db	813	ILAYKLIVQTHGVNDVPDKELVLTNRNVNSDGIINQRAFTNYLSAW-----ATNASSTEL	---	---	868
Call 59, 751-765, 1989				QY	1029	LAA--YSTSVNLTSQGVLASRFMAYHKPLKNSQD-----NEYDLKIPKSLPLVYAQMPPYLGLTDT	---	---	1070
A:Title: The Drosophila patched gene encodes a putative membrane protein required for se				Db	869	LRANCIRNRANGASQGLYPEPROYFHQP--NEYDLKIPKSLPLVYAQMPPYLGLTDT	---	---	926
A:Reference number: A33468; MUID:90058658; PMID:2582494				QY	1071	ELAAINITADLRKVPGTDPAFV--FPYTTTNVFEQYLTILPEGLFMLSCLVFTFAVSC	---	---	1128
A:Accession: A33468				Db	927	QI-KTLIGHIRDLGVKVEGFLPNYPGIPPFIFWEQYMT-LRSSLAMILACVLLAALVIV	---	---	984
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra				QY	1129	LLGLDLRSLGNLLSIVMILVDTVGFMALWDISYNAVSLINLVSAVGMSVEFVSHITRS	---	---	1188
A:Molecule type: mRNA				Db	985	SLLSVMAAVLVILSVLSLAQAQIFGATLLGKLSAIPAVILLISVGMCLCF--NVLLIS	---	---	1042
A:Residues: 1-110,'R',112-273,'G',275-331,'R',333-635,'P',637-861,'DVF',878,'Y',880-1299				QY	1189	FAISTKPTLWRAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRMLLTITL	---	---	1248
A:Cross-references: UNIPARC:UPI000016BCA3; GB:M28418; GB:M28999; MID:g552097; PID:g55209				Db	1043	LGFMITSVGNRRQRRVQLSQMSLGPLVHGMLTSGVAVFMLSTSPPEFVIRHFECWLLLVLC	---	---	1102
C:Genetics:				QY	1249	LGLHGLVFLPVILSYGPDVNPALALOKRAEAAVAVMWASCNHPSRVSTADNIYVN	---	---	1308
A:Gene: FlyBase:ptc				Db	1103	VGACNSLLVFPILLSVMVGPEAB-LVPLE-----HPDRISTPSPLPVR	---	---	1143
A:Cross-references: FlyBase:FBgn0003892				QY	1309	HS	1310		
A:Map position: 2 44D3-D4				Db	1144	SS	1145		
C:Superfamily: Drosophila membrane protein patched									
C:Keywords: glycoprotein; transmembrane #status predicted <TM01>									
F;74-92/Domain: transmembrane #status predicted <TM02>									
F;427-448/Domain: transmembrane #status predicted <TM03>									
F;456-503/Domain: transmembrane #status predicted <TM04>									
F;529-555/Domain: transmembrane #status predicted <TM05>									
F;557-585/Domain: transmembrane #status predicted <TM06>									
F;677-699/Domain: transmembrane #status predicted <TM07>									
F;967-1017/Domain: transmembrane #status predicted <TM08>									
F;1019-1047/Domain: transmembrane #status predicted <TM09>									
F;1061-1086/Domain: transmembrane #status predicted <TM10>									
F;1093-1121/Domain: transmembrane #status predicted <TM10>									
F;142,298,335,388,807,861,1194,1271/Binding site: carbohydrate (Aen) (covalent) #status									
Query Match									
Best Local Similarity									
Matches									
321									
43									
381									
103									
437									
154									
487									
205									

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE IS BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:13:14 ; Search time 265 Seconds  
(without alignments)  
3546.278 Million cell updates/sec

Title: US-10-736-769-4  
Perfect score: 6909  
Sequence: 1 MAEAGURGMLLALLRLAQ.....GSIKAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	6872.5	99.5	1359	1 NPC1L1_HUMAN	Q9uhc9 homo sapien
2	5421.5	78.5	1331	1 NPC1L1_RAT	Q6t3u3 rattus norv
3	5407	78.3	1333	1 NPC1L1_MOUSE	Q6t3u4 mus musculu
4	3179	46.0	1132	2 Q4T749_TETNG	Q4t749 tetraodon n
5	2413	34.9	1277	2 Q9JLG3_CRIGR	Q9jlg3 cricetulus
6	2412	34.9	1277	1 NPC1L1_PIG	P56941 sus scrofa
7	2405	34.8	1276	2 Q8MKD8_FELCA	Q8mkd8 felis silve
8	2404.5	34.8	1276	2 Q9GK52_CANFA	Q9gk52 canis fami
9	2402.5	34.8	1278	1 NPC1L1_HUMAN	O15118 homo sapien
10	2402	34.8	1289	2 Q59GR1_HUMAN	Q59gr1 homo sapien
11	2400	34.7	1277	2 Q7TMD4_MOUSE	Q7tmd4 mus musculu
12	2399	34.7	1276	2 Q9NQ0Q_FELCA	Q9nq0q felis silve
13	2395	34.7	1276	2 Q8M149_FELCA	Q8m149 felis silve
14	2385	34.5	1278	1 NPC1L1_MOUSE	O35604 mus musculu
15	2367.5	34.3	1277	2 Q9GLC9_BOVIN	Q9glc9 bos taurus
16	2361	34.2	1286	2 Q9TT75_RABIT	Q9tt75 corytolagus
17	2338	33.8	1209	2 Q4RWY5_TETNG	Q4rwy5 tetraodon n
18	2282	32.7	1287	2 Q9VL24_DROME	Q9vl24 drosophila
19	2256	32.7	1287	2 Q9U5W1_DROME	Q9u5w1 drosophila
20	2253	31.6	1287	2 Q7YU59_DROME	Q7yu59 drosophila
21	2205	31.9	1291	2 Q7Q409_ANOGA	Q7q409 anopheles g
22	1864.5	27.0	1223	2 Q9VRC9_DROME	Q9vrc9 drosophila
23	1760	25.5	1003	2 Q7PS03_ANOGA	Q7ps03 anopheles g
24	1725	25.0	1275	2 Q5LNK7_MAGGR	Q5lnk7 magnaporthe
25	1712.5	24.6	1361	2 Q7XU07_ORYSA	Q7xub7 oryza sativ
26	1700.5	24.6	1271	2 Q5BBG1_EMENI	Q5bbg1 aspergillus
27	1659.5	24.0	1330	2 Q5KGS9_CRYNE	Q5kgs9 cryptococcu
28	1658.5	24.0	1330	2 Q5SSD4_CRYNE	Q5ssd4 cryptococcu
29	1643	23.8	1273	2 Q4WNG5_ASFFU	Q4wns5 aspergillus
30	1614.5	22.4	1295	2 Q410K4_GIBZE	Q410k4 gibberella
31	1549	22.4	1264	2 Q6BT03_DEBHA	Q6bt03 debaryomyce

32	1527.5	22.1	1275	2 Q9SHN9_ARATH	Q9shn9 arabidopsis
33	1511.5	21.9	1342	2 Q9TVK6_DICDI	Q9tvk6 dictyosteli
34	1508	21.8	1239	2 Q6CBAL_YARLI	Q6cbal yarrowia li
35	1500	21.7	1162	2 Q7RWL9_NEUCR	Q7rw19 neurospora
36	1461.5	21.2	1256	2 Q59ZV0_CANAL	Q59zv0 candida alb
37	1452.5	21.0	1397	2 Q5S1C5_DICDI	Q5s1c5 dictyostell
38	1437.5	20.8	1489	2 Q4PEB3_USTMA	Q4peb3 ustilago ma
39	1394	20.2	1055	2 Q9SVF0_ARATH	Q9svf0 arabidopsis
40	1386.5	20.1	620	2 Q4H344_CIOIN	Q4h344 ciona intes
41	1359	19.7	1382	2 Q618V4_CAEBR	Q618v4 caenorhabdi
42	1341.5	19.4	1383	1 NPC1L1_CAEEL	Q19127 caenorhabdi
43	1329	19.2	1170	2 Q12200_YEAST	Q12200 saccharomyc
44	1323	19.1	1339	2 Q50RB6_ENTHI	Q50rb6 entamoeba h
45	1312.5	19.0	1178	2 Q750G1_ASHGO	Q750g1 ashbya goss

ALIGNMENTS

RESULT 1	
NPCL1_HUMAN	
ID	NPCL1_HUMAN STANDARD; PRT; 1359 AA.
AC	Q9UHC9; Q6R3Q4; Q9UHC8;
DT	13-SEP-2005 (Rel. 48, Created)
DT	13-SEP-2005 (Rel. 48, Last sequence update)
DT	13-SEP-2005 (Rel. 48, Last annotation update)
DE	Niemann-Pick C1-like protein 1 precursor.
GN	Name=NPCL1;
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo.
OX	NCBI_TaxID=9606;
RN	[1]_TaxID=9606;
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2 AND 3), AND TISSUE SPECIFICITY.
RP	MEDLINE=20247253; PubMed=10783261; DOI=10.1006/geno.2000.6151;
RX	Davies J.P., Levy B., Ioannou Y.A.;
RA	"Evidence for a Niemann-Pick C (NPC) gene family: identification and characterization of NPC1L1.";
RT	Genomics 65:137-145(2000).
RL	[2]
RN	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), AND TISSUE SPECIFICITY.
RP	PubMed=14976318; DOI=10.1126/science.1093131;
RP	Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M., Wang L.,
RA	Tetloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
RA	Murgolo N., Graziano M.P.;
RT	"Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption.";
RL	Science 303:1201-1204(2004).
RN	[3]
RP	TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RP	PubMed=15671032; DOI=10.1074/jbc.M409110200;
RA	Davies J.P., Scott C., Olsh K., Liapis A., Ioannou Y.A.;
RA	"Inactivation of NPC1L1 causes multiple lipid transport defects and protects against diet-induced hypercholesterolemia.";
RL	J. Biol. Chem. 280:12710-12720(2005).
RN	[4]
RP	INDUCTION.
RP	PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;
RX	van der Veen J.N., Kruit J.K., Havinga R., Baller J.F.W., Chimini G.,
RA	Lestavel S., Staels B., Groot P.H.E., Groen A.K., Kuipers F.;
RA	"Reduced cholesterol absorption upon PPARdelta activation coincides with decreased intestinal expression of NPC1L1.";
RL	J. Lipid Res. 46:526-534(2005).
RN	[5]
RP	FUNCTION.
RP	PubMed=15928087; DOI=10.1073/pnas.0500269102;
RX	Garcia-Calvo M., Lienock J., Bull H.G., Hawes B.E., Burnett D.A.,
RA	Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Demers P.A.,
RA	Graziano M.P., Hughes M., Macintyre D.E., Ogawa A., O'Neill K.A.,
RA	Iyer S.P.N., Shevell D.B., Smith M.M., Tang Y.S., Makarewicz A.M.,
RA	Ujjainwalla F., Altmann S.W., Chapman K.T., Thornberry N.A.;



Db 301 LLVGRVAPARDKSKWDPKKGYSLSKUSFSFTHLLGQFFQGWGTWASWPTILVLSV 360  
QY 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFFRTNOVILTAPNRSSV 420  
Db 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFFRTNQVILTAPNRSSV 420  
QY 421 RYDSLLGPKNFGSGILDLDLLELLELQERLRHLQWSPQAQRNLSIQDICYAPLNPDMT 480  
Db 421 RYDSLLGPKNFGSGILDLDLLELLELQERLRHLQWSPQAQRNLSIQDICYAPLNPDMT 480  
QY 481 SLYDCCINSLLQVFQNNRTLLLTANQTLWGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540  
Db 481 SLYDCCINSLLQVFQNNRTLLLTANQTLWGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540  
QY 541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
Db 541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
QY 601 RAFQRMAGMFQVTFTAERSLEDEINRTTAEDLPFPATSYVIFLYISLALGYSYSSWSRV 660  
Db 601 RAFQRMAGMFQVTFTAERSLEDEINRTTAEDLPFPATSYVIFLYISLALGYSYSSWSRV 660  
QY 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720  
QY 721 YQRLPRPEPEPREVHIGRALGRVAPSMILCSSEALCFPLGALTMPAVRTTALTSGLAV 780  
Db 721 YQRLPRPEPEPREVHIGRALGRVAPSMILCSSEALCFPLGALTMPAVRTTALTSGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCVKPOELPPQGGELLIGFFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCVKPOELPPQGGELLIGFFQKAYAPFL 840  
QY 841 LHWITRGVLLFLALFGVLSYSMCHISVGLQDELALPKDSYLLDYFLFNRYFVGAPV 900  
Db 841 LHWITRGVLLFLALFGVLSYSMCHISVGLQDELALPKDSYLLDYFLFNRYFVGAPV 900  
QY 901 YFVTTILGVNFSSEAGNNAICSSAGCENNFSFTQKIQVATFPPQSQYLAIIPASSWVDFFDM 960  
Db 901 YFVTTILGVNFSSEAGNNAICSSAGCENNFSFTQKIQVATFPPQSQYLAIIPASSWVDFFDM 960  
QY 961 LTPSSCCRLYISGPNKDKPCPSTVNSLNCNKCMSITMGSVRPSVEQFHKLWFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDKPCPSTVNSLNCNKCMSITMGSVRPSVEQFHKLWFLNDRP 1020  
QY 1021 NIKCPKGLAAYSTSVNLTSQGVLT-----ASRFMAYH 1053  
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLT-----ASRFMAYH 1053  
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFVEQYLTILPEGL 1113  
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFVEQYLTILPEGL 1113  
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTGFMALWDISYNAVSINLVS 1173  
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTGFMALWDISYNAVSINLVS 1200  
QY 1174 AVGMSVEFVSHITRSFAISTKTWLERAKEATISMGSAVFAVAMTNLPGLILVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITRSFAISTKTWLERAKEATISMGSAVFAVAMTNLPGLILVLGLAKAQ 1260  
QY 1234 LIQIFFRNLNLTLLGLLHGLVFLPVILSVYVGPDPNPALEOKRAEAAVAVMVASCP 1293  
Db 1261 LIQIFFRNLNLTLLGLLHGLVFLPVILSVYVGPDPNPALEOKRAEAAVAVMVASCP 1320  
QY 1294 NHPRSVSTADNIYVNHSPFGSKGAGAIISNPLPNNGRQF 1332  
Db 1321 NHPRSVSTADNIYVNHSPFGSKGAGAIISNPLPNNGRQF 1359

RESULT 2

NPCL1 RAT  
ID NPCL1 RAT STANDARD; PRT; 1331 AA.  
AC O6T3U3;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Niemann-Pick Cl-like protein 1 precursor.  
GN Name=Npc1l1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND SUBCELLULAR  
RP LOCATION.  
RC STRAIN=Sprague-Dawley;  
RX PubMed=14976338; DOI=10.1126/science.1093131;  
RA Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,  
Tetzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,  
Murgolo N., Graziano M.P.;  
RA "Niemann-Pick Cl like 1 protein is critical for intestinal cholesterol  
absorption.";  
RL Science 303:1201-1204(2004).  
RN [2]  
RP SUBCELLULAR LOCATION, AND CHARACTERIZATION.  
RX PubMed=1577641; DOI=10.1016/j.bbagen.2004.12.021;  
RA Iyer S.P.N., Yao X., Crona J.H., Hoos L.M., Tetzloff G.,  
Davis H.R. Jr., Graziano M.P., Altman S.W.;  
RA "Characterization of the putative native and recombinant rat sterol  
transporter Niemann-Pick Cl Like 1 (NPCL1) protein.";  
RL Biochim. Biophys. Acta 1722:282-292(2005).  
RN [3]  
RP FUNCTION.  
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;  
RA Garcia-Calvo M., Lisnock J., Bull H.G., Hawn B.E., Burnett D.A.,  
Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Demers P.A.,  
Graziano M.P., Hughes M., MacIntyre D.E., Ogawa A., O'Neill K.A.,  
Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,  
Ujjainwalla F., Altman S.W., Chapman K.T., Thornberry N.A.;  
RA "The target of ezetimibe is Niemann-Pick Cl-Like 1 (NPCL1).";  
RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137(2005).  
CC -!- FUNCTION: Play a major role in cholesterol homeostasis. Is  
critical for the uptake of cholesterol across the plasma membrane  
of the intestinal enterocyte. Is the direct molecular target of  
ezetimibe, a drug that inhibits cholesterol absorption (by  
similarity). Lack of activity leads to multiple lipid transport  
defects. The protein may have a function in the transport of  
multiple lipids and their homeostasis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation  
of brush border membranes from proximal enterocytes suggests  
considerable association with the apical membrane fraction. Exists  
as a predominantly cell surface membrane expressed protein.  
CC -!- TISSUE SPECIFICITY: Small intestine showed the highest level of  
expression. Expression in other tissue including gall bladder,  
liver, testis, and stomach is also observed. Along the duodenum-  
ileum axis, the levels vary in different segments of intestine,  
with peak expression in the proximal jejunum. Protein expression  
is confined to the enterocyte. Discrete localization to the  
epithelial layer bordering the luminal space along the crypt-  
villus axis. Protein expression in the enterocyte is observed  
closest to the luminal space. Expression in enterocytes from the  
proximal (jejunum) but not in the distal (ileum) region.  
CC -!- PTM: Highly glycosylated.  
CC -!- SIMILARITY: Belongs to the patched family.  
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use as long as its content is in no way modified and this statement is not  
removed.  
CC -----  
CC

Mon Apr 10 07:16:16 2006

DR EMBL; AY437867; AAR97888.1; -, mRNA.  
 DR RGD; 130135; Npc11.  
 DR InterPro; IPR004765; NP\_C type.  
 DR InterPro; IPR003392; PaTched.  
 DR InterPro; IPR000731; SSD\_STM.  
 DR Pfam; PF02460; Patched; 1.  
 DR TIGRFAMs; TIGR00917; 2A060601; 1.  
 DR PROSITE; PS0156; SSD; 1.  
 KW Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;  
 KW Steroid metabolism; Transmembrane.  
 DR SIGNAL 1 20 Potential.  
 DR CHAIN 21 1331 Niemann-Pick C1-like protein 1.  
 DR TOPO\_DOM 21 282 Extracellular (Potential).  
 DR TRANSMEM 283 303 1 (Potential).  
 DR TOPO\_DOM 304 352 Cytoplasmic (Potential).  
 DR TRANSMEM 353 373 2 (Potential).  
 DR TOPO\_DOM 374 632 Extracellular (Potential).  
 DR TRANSMEM 633 653 3 (Potential).  
 DR TOPO\_DOM 654 665 Cytoplasmic (Potential).  
 DR TRANSMEM 666 686 4 (Potential).  
 DR TOPO\_DOM 687 696 Extracellular (Potential).  
 DR TRANSMEM 697 717 5 (Potential).  
 DR TOPO\_DOM 718 742 Cytoplasmic (Potential).  
 DR TRANSMEM 743 763 6 (Potential).  
 DR TOPO\_DOM 764 776 Extracellular (Potential).  
 DR TRANSMEM 777 797 7 (Potential).  
 DR TOPO\_DOM 798 846 Cytoplasmic (Potential).  
 DR TRANSMEM 847 867 Extracellular (Potential).  
 DR TOPO\_DOM 868 1113 9 (Potential).  
 DR TRANSMEM 1114 1134 Cytoplasmic (Potential).  
 DR TOPO\_DOM 1135 1142 10 (Potential).  
 DR TRANSMEM 1143 1163 Extracellular (Potential).  
 DR TOPO\_DOM 1164 1165 11 (Potential).  
 DR TRANSMEM 1166 1186 Cytoplasmic (Potential).  
 DR TOPO\_DOM 1187 1206 Extracellular (Potential).  
 DR TRANSMEM 1207 1227 12 (Potential).  
 DR TOPO\_DOM 1228 1242 13 (Potential).  
 DR TRANSMEM 1243 1263 Cytoplasmic (Potential).  
 DR TOPO\_DOM 1264 1331 SSD.  
 DR DOMAIN 632 797 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 53 53 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 85 85 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 138 138 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 244 244 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 416 416 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 431 431 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 464 464 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 479 479 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 497 497 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 506 506 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 606 606 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 626 626 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 909 909 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 917 917 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 996 996 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 1038 1038 N-linked (GlcNAc. .) (Potential).  
 DR CARBOHYD 1076 1076 N-linked (GlcNAc. .) (Potential).  
 DR SEQUENCE 1331 AA; 146415 MW; 2E10BF2E3A337F70 CRC64;  
 Query Match 78.5%; Score 5421.5; DB 1; Length 1331;  
 Best Local Similarity 77.8%; Pred. No. 0;  
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;  
 QY 1 MAEAGLRGMLLWALLRLAQSEPYTTIHQGYCAFYDEGCKNPGLSGSLMTLSNVSCLSN 60  
 DB 1 MAEAWL-GMLLWALLLSAAQGLYTPKHEAGVCTFYEECKNPGLSGSLTSLNSVSCLSN 59  
 QY 61 TPARKITGDHLLILQKICPLRYTGPNTQ-ACCBAKOLVSLBASLSITKALLTRCPACSDN 119  
 DB 60 TPARHVTGHEALLQRICPLRYNPNTTFACCSTKQLLSLESSMSITKALLTRCPACSDN 119  
 QY 120 FVNLHCHNTCSPNQSLFPIVTRVAQLGAGOLPAVAYEAFYQHSFBAQSDCSRVVRVPA 179  
 DB 120 FVNLHCHNTCSPDQSLFINVTRVVERGAGEPPAVVAYEAFYQHSFBAQSDCSRVRI 179

120 FVSLHCHNTCSPDQSLFINVTRVVERGAGEPPAVVAYEAFYQHSFBAQSDCSRVRI 179  
 QY 180 AATLAVGTMCVGSALCNAORWLNFGQDTGNGLAPLDITPHLEPQGAQSGIQLNEG 239  
 DB 180 AASLAVGSMCGVGSALCNAORWLNFGQDTGNGLAPLDITPHLEPQGAQSGIQLNEG 239  
 QY 240 VARNESQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILCSFVAVT 299  
 DB 240 IAPCNESQGDSDAVCSQDCAASCFCVIPPPEALRPSFYMGMPGMLALIIITFAVFLIS 299  
 QY 300 ILLVGRVAPARDKSMVDPKKGTSLSDKLSPTHTLLGQFQGHGTWVASPLTLVLIS 359  
 DB 300 AVLRLRVVSNRKNKAEGPQAPKLPHKHLSPHTILGRFFQNNGTTRVASPLTLVLIS 359  
 QY 360 VTPVVALAAGLVFTETLTPDVELMSAPNSQARSEKAFHQHFGPFPRTNQVLTAPNRSS 419  
 DB 360 FTWIALAAGLTFIELTTPDVELMSAPNSQARSEKAFHQHFGPFPRTNQVLTAPNRSS 419  
 QY 420 YRYDSLILGPKNFGSGLDLDLLELELEQERLRLHQVMSPEAQRNISLQDICVAPLNPN 479  
 DB 420 KYKDSLILGSKNFGSGLDLDLLELELEQERLRLHQVMSPEAQRNISLQDICVAPLNPN 479  
 QY 480 TSLYDCINSLLOYPONNRTLLILTRANQTLMGTSQVDMKDHFLYCANAPLTFKQGTALA 539  
 DB 480 TSLSDCCVNSLLQYFQNNRTLLMLTANQTLNGQTSVLDWKDHFLYCANAPLTFKQGTALA 539  
 QY 540 LSCMADYGAPVPFPLAIGVKGKDYSEAEALIMTFSINNYPAGDPRLQAQKWEAEFLBE 599  
 DB 540 LSCMADYGAPVPFPLAIGVKGKDYSEAEALIMTFSINNYPAGDPRLQAQKWEAEFLBE 599  
 QY 600 MRAPQRMAGMFQVTFEATERSLEDEINRTAEDLPATSYIVIFYISLALGSYSWSR 659  
 DB 600 MESFQNTSDKQVAFSAERSLEDEINRTIQDLPVFAVSIIYFVLYISLALGSYSRCSR 659  
 QY 660 VMVDSKATIGLGVAVVLGAVMAAGFFSVLGRSSLVILQVVPFVLVSGADNIFIVL 719  
 DB 660 VAVESKATIGLGVAVVLGAVMAAGFFSVLGRSSLVILQVVPFVLVSGADNIFIVL 719  
 QY 720 EYQRLPRRPEPREVHIGRALGRVAPSMILCSLSEACFFLGLALTMPAVRTFALTSGLA 779  
 DB 720 EYQRLPRRPEPREVHIGRALGRVAPSMILCSLSEACFFLGLALTMPAVRTFALTSGLA 779  
 QY 780 VILDFLQMSAFVALLSDSKRQESRLDVCCVCKPOELPPQGGEGLLIGFQKAYPF 839  
 DB 780 IILDFLQMTAFVALLSDSKRQESRPDLVCCFSTRKLPPEKEGELLRFPRKIYAPF 839  
 QY 840 LLHWITRGVLLLPALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLNRYFVEYGP 899  
 DB 840 LLHRTFIRVVMVLLFTLFGANLYLMCNINVLGDQELALPKDSYLLDYFLNRYLEVGP 899  
 QY 900 VYFVTTLYNFPSSBAGNNAICSSAGCNNSFTQKIQYATEFPEQSYLAI PASSWDDFID 959  
 DB 900 VYFVTTSGFNFSSBAGNNAICSSAGCKSFSLTQKIYASEPFDQSYVAIAASSWDDFID 959  
 QY 960 WLTP-SSCCRLYISGPNKDFCPTVNSLNCNKMSTMGSVPSVPSVEQFKYLPWLND 1018  
 DB 960 WLTPSSSSCCRLYIRGPHKDFCPTDTSFNCLKCNMRTLGPVRPTAEQFKYLPWLND 1019  
 QY 1019 RPNKCPKGGAAAYSTSVNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELAANITA 1078  
 DB 1020 PPNIRCPKGGAAAYRSTSVNLSSDQGVITASQFMAVHKPLRNSQDYTEALRAARELAANITA 1079  
 QY 1079 DLKRVPGDPAPEFPFPTITNVFVEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSG 1138  
 DB 1080 DLKRVPGDPAPEFPFPTITNVFVQYLTILPEGLIFTLALCFVPTFVVVYLLGLDMSG 1139  
 QY 1139 LNLLSIYMLIVDTVGFVWALWDISVNAVSLINLSVAGMSVEFVSHITRESFAISKPTWL 1198  
 DB 1140 LNLLSIIMILVDITGLMAVMSISVNAVSLINLSVAGMSVEFVSHITRESFAVSKPTWL 1199  
 QY 1199 ERAKEATISMSAPVAVMTNLPGLVLGLAKAQLIOIFFRLNLLITLGLLHGLVFL 1258  
 DB 1200 ERAKDATVPMGSAPVAVMTNFPGLILGLFAQAQLIOIFFRLNLLITLGLLHGLVFL 1259



Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRCWLLWALLLRLAQSEBPTTTHQPGYCAFYDECGKNPGLSGMLTSLNSVCLSNTPA	63
Db	3	AAWQGLLWALLLRLAQSEBPTTTHQAGCTTTECCGNPGLSGMLTSLNSVCLSNTPA	62
Qy	64	RKTGDLHLLQKTCPLRYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSDNFVN	122
Db	63	RHTGDLHALLQRCPLRYNGPNDTYACCSKQLVSLDSSLITKALLTRCPACSENFS	122
Qy	123	LHCHNTCSPNQSFLNTRVQAOLGAGQLPAVVAEAFYQHSFAQSDSCSRVUPAAAT	182
Db	123	IHCHNTCSPDQSLFNTRVQVDPGQLPAVVAEAFYQHSFAQSDSCSRVUPAAAS	182
Qy	183	LAVGTCMCGYSGALCNARWLNFGDGTGNLGLAPDITPHLLPQGAQVSGIQPLNEGVAR	242
Db	183	LAVGTCMCGYSGALCNARWLNFGDGTGNLGLAPDITPHLLPQGAQVSGIQPLNEGVAR	242
Qy	243	CNESQDDVATCSCQDCAASCATARPQALDSTFYLGMPGSLVLIILCSFVAVTILL	302
Db	243	CNESQDDVATCSCQDCAASCATARPQALDSTFYLGMPGSLVLIILCSFVAVTILL	302
Qy	303	VGRVAPADKSMVDPKGTSLSKLSFSHTLLGQFQCGWGTWVASMPLTILVLSVIP	362
Db	303	VYLRVASNRKNKTAGSQEAPNLPKRFRSPHTVLGRFESNGTTRVASMPLTILVLSVIP	362
Qy	363	VVALAAGLVFTLTTDPVELWSPNSQARSEKAFHQHFGPPFRTNQVILTPAPNSRYR	422
Db	363	VVALSGLVFTLTTDPVELWSPNSQARSEKAFHQHFGPPFRTNQVILTPAPNSRYR	422
Qy	423	DSLLGPKNFSGILLDLLELLELQERLRLHQLVMSPEARQNSISLQDIQYAPLNDTSL	482
Db	423	DSLLGPKNFSGILLDLLELLELQERLRLHQLVMSPEARQNSISLQDIQYAPLNDTSL	482
Qy	483	YDCCINSLLQYFQNNRTLLLTANQTLMGOTSQVDWKHFLCANAPLTKDGTALALSC	542
Db	483	TDCCVNSLLQYFQNNRTLLLTANQTLMGOTSLVDWKHFLCANAPLTKDGTALALSC	542
Qy	543	MADYGAPVFPFLAIGYKGYKGYSEAEALIMTSLNNYPAGDPRLAQALWEAFLEWRA	602
Db	543	IADYGAPVFPFLAIGYKGYKGYSEAEALIMTSLNNYPAGDPRMAHAKWEAFLEWRA	602
Qy	603	FORMAGMFOVPTAERSLEDEINRTAEDLPFATSYIVILYISLAGSYSSRSRVAV	662
Db	603	FORSTADKFGIAFSAERSLEDEINRTAEDLPFATSYIVILYISLAGSYSSRSRVAV	662
Qy	663	DSKATIGLGVAVVGLGVAAMGFFSVGLRSLVLQVVPFLVLSVGADNIFIFVLEYQ	722
Db	663	DSKATIGLGVAVVGLGVAAMGFFSVGLRSLVLQVVPFLVLSVGADNIFIFVLEYQ	722
Qy	723	RLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTMPMPAVRTFALTSLAVIL	782
Db	723	RLPRMPEQREAHIGRTLSVAPSMLLCSLSEAICFFLGALTSMMPAVRTFALTSLGLIIF	782
Qy	783	DFLLQMSAFVALLSLDSKROEASRLDCCVCKQDELPPCGGSLILGPFQKAYAPPELLH	842
Db	783	DFLLQMTAFVALLSLDSKROEASRPDVCCFSSRNLPKPKQGGLLCCFFRKLYTFELLH	842
Qy	843	WITRGVLLFLALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLPLNRYFVGPAPVYF	902
Db	843	RFITRPVLLFLVLFGANLYLNCNISVGLDQDALPKDSYLLDYFLPLNRYLVEGPEVYF	902
Qy	903	VTTLLGYNFSFAGMNAICSSAGCNFSTQKIQVATFEPQSYLAIPASSWVDFTDMLT	962
Db	903	DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIQVSEFPNQSYVIAAASSWVDFTDMLT	962
Qy	963	P-SSCCRIYISGPNKDFCPTVNSLNCNMSITWGSVRPSVEQFHKYLPLNDRPN	1021
Db	963	PSSSCRIYTRGPHKDFCPTDTSFNCLKNCMRTLGPVPTTEQFHKYLPLNDRPN	1022
Qy	1022	IKCPKGLGAAYSTSVNLTSQVLAASFMAVYHKLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGLGAAVTSVNLSSDQGLIASQFMAVYHKLKNSQDYTEALRASRLAANITAEAR	1082

1082 KVPCTDPAFEVPPYTTITNVFEYOYLITLPEGLFMLSCLVPTFAVSCILGLDLRSGLN 1141

1083 KVPCTDPAFEVPPYTTITNVFEYOYLITLPEGLFMLSCLVPTFAVSCILGLDLRSGLN 1142

1142 LLSIVMILVDITVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSFAISKPTWLER 1201

1143 LLSIVMILVDITVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSFAISKPTWLER 1202

1202 KEATISMGSAVAGVAMTNLPGIILVILGLAKAQIIOIFFERLNLITLLGLLHGLVFLPY 1261

1203 KDATIFMGSAVAGVAMTNLPGIILVILGLAKAQIIOIFFERLNLITLLGLLHGLVFLPY 1262

1262 LSVGVDPNVPALAEOKRAEEAAVAAVMVASCNHPHSRVSTADNIYVNHVSFEQS - IKGAGA 1320

1263 LSVGVDPNVPALAEOKRAEEAAVAAVMVASCNHPHSRVSTADNIYVNHVSFEQS - IKGAGA 1321

1321 ISNFPENNGRQF 1332

1322 ASSSLPKSDQKF 1333

RESULT 4

QAT749 TETNG

ID QAT749 TETNG PRELIMINARY; PRT: 1132 AA.

AC QAT749;

DT 13-SEP-2005 (T-EMBLrel. 31, Created)

DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)

DE Chromosome undetermined SCAF8317, whole genome shotgun sequence. (Fragment)

DE ORFNames=GSTENG0005930001;

GN Tetraodon nigroviridis (Green puffer).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

OC NCBI\_TaxID=99883;

EN [1]

RP NUCLEOTIDE SEQUENCE.

RA Jallou O., Aury J.M., Brunet P., Petit J.L., Stange-Thomann N., Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B., Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S., Anhouard V., Jubin C., Castelli V., Catolico L., Poulain J., De Berardinis V., Biemont C., Skalli Z., Brottier P., Coutanceau J.P., Gouzy J., Craud C., Duprat S., Brottier P., McKernan K.J., McEwan P., Bosak S., Parra G., Lardier G., Chapple C., Guigo R., Zody M.C., Mesirov J., Kellis M., Volff J.N., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Lindblad-Toh K., Birren B., Saurin W., Scarpelli C., Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C., Wincker P., Lander E.S., Weissbach J., Roest Crolius H.; "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.";

RT Nature 431:946-957(2004).

RL Nature 431:946-957(2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research; Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

RL -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CAAB01008317; CAF91283.1; -; Genomic\_DNA.

DR NON TER 1

FT NON TER 1

SQ SEQUENCE 1132 AA; 125923 MW; 4283CF4CA0BBDE97 CRC64;

Query Match 46.0%; Score 3179; DB 2; Length 1132;

Best local similarity 52.0%; Pred. No. 4.3e-216;

Matches 641; Conservative 167; Mismatches 264; Indels 160; Gaps 11;

Qy 87 TOACCSAKQLVSLASLSITKALLTRCPACSDNFVNHLCHNTCSPNQSFLNTRVQAOL- 145

Db 1 TYACCSINQLSSLEASLSLASKAVLIRCPSCADNFHFCITTCSPQOQTKTKVTKVNTV 60





Db 4 RGAFLGLLLLLCPVQFSQ-----SCVWTGEGC-----IASGD-----KRYNCRVSGPPKPL 51

Qy 67 TGDHLILLOKICRLTYGTGNTACCSAKQIVLSLEASISITKALLITRCPACSDNFVNLHCH 126

Db 52 PEDGYDLVQELCPGFFPG-NVSLCCDVQQLRTLDKNIQLPLQFLSRCPSCFYNLMLNLFCE 110

Qy 127 NTCSPNQSILFINVTR-----VAQLGAGQLPAVAYEAFYQHSFPAEQSDSCSRVRVPAAT 182

Db 111 LTCSPRQSQFLNVTATEDYVDPVTNQTKNVKELEYVGETFANAMYNACRDVEAPSSNE 170

Qy 183 LAVGTWCVGYGSALCNAQRWLNFGQDTGNGLAPLDT--PHLEPGQAVSGIQPLNEGV 240

Db 171 KALGLLCGREAQA-CHNATNWIETFNKNDGQAPFTTIPFSDLPT-----HGMPEMNNAT 224

Qy 241 ARNESQGDVATCSQDCQAASC-----PAIARPOALDSTFYLGQMPGSLVLIIL 291

Db 225 KGDESVDVETGFCSCQDCSIVCGPKPQPPPPVPMRILGLDAMYIMSSYNMAFLIVFP 284

Qy 292 CSVPANVTILLVGRVAPARDKSKMVDPKGTSLSKLSPTHTL----- 336

Db 285 GAFFAVWCY-----RKRYFVSEYTPIDGNIAPSVNSDKQAFCFCDPLCAAF 331

Qy 337 ---LGQFFQCGWGTWASWPLTILVLSVIPVALAAGLVFTELTDPVELWSAPNSQARSE 393

Db 332 ERGLRLFAQWGAFCVRHPCGVFFSLAFIVACSSGLVIRVTTDPVDLWSAPGSQARRE 391

Qy 394 KAFHDQHFQFFFTNQTAVPNSRYRSDLSLLGPK-NFSGILDLDLLELELELERLR 452

Db 392 KEYFDTHFGFFEMEQLIIRATNNQSHIYHPYPAGADVPGPPLSRDLHLQVLDLQTATE 451

Qy 453 HLOVWSEPAQRNLSLODICVAPLNPNQTSYDCINSLLOQYFQNNRTLLLTANQTMGQ 512

Db 452 N--ITASYNNEVTLODICLAPLSYNN--NCTILSVLNYFQNSHVL----- 495

Qy 513 TSOV-----DWKDHLYCANAPITFKDGTALALSCHADYCAPVPPFLATGCGYKDY 564

Db 496 DHQVGDFFFVYADYHTHLYCYVRAPASLNDASLLHDPCLGTGFGPFPFVLVLGVLGQDQY 555

Qy 565 SEAEALIMFTSLNNYPAGPRLAQAKLWEEAFLEENRAFORRMAGMFQVTFTAERSLDE 624

Db 556 NNATALVITFPVNNYNDTEKLOQAQWSEFINFVNKYN---PNLTISFPAERIEDE 612

Qy 625 INRTAEDLPIFATSYVIVFLYSLALGSYSMSRVMSDKATFLGGVAVVLGAVMAAM 684

Db 613 LNRESNLDLTILISVAINFLYSLIALGHKSCRLVDSKISGIAGILIVLSSVACSL 672

Qy 685 GFPSYLGIRSSVLQVDFVLVSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVA 744

Db 673 GIPSYIGVPLTLIVIEVIFPLVAVGVDNIFILVQTYQDERLQGETLQQLGRVLGEVA 732

Qy 745 PSMILCLSEAICFFPLGALTPMAVETPALTSLGLAVILDFLQMSAFVALLSDSKRQA 804

Db 733 PSWFLSFSETVAFGLGSLVVPVAVTFFSLFAGNAVILDFLQITCFVSLGLLDIKRQEK 792

Qy 805 SRLDVCCKVQBELPPPGQ-----EGLLLGFFOKAYAPFLLHWITRGVLLFLALFGVS 860

Db 793 NRLDVVCCVQGAE--DGAGVQASECLFRFPKNSYAPILLKDWMPRIAVFVGVLSFS 849

Qy 861 LYSMCHISVGLDQELAPKDSYLLDYLFLNRYFEVGAFFVFTTLYGFFSSAGHNAIC 920

Db 850 IAVLNKVEIGLQSLSMPPDSYMYDYFQSLSRYLHAGPPVYFVSEGHNTYSLKGNMYC 909

Qy 921 SSAGCNFSPQIKQYATFPEQSYLAIPASSWVDDPIDWLTP-SSCCRLYISGPNKDF 979

Db 910 GGLGCNNDLSUQIIFTAQLDNVTIRIGFAPSSWIDDYFDWIKRQSSCCRYNS---TDQF 966

Qy 980 CPSTVNSLNLKNCMSITW-GSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNL 1038

Db 967 CNASVVDPTCIR-CRPLTSEGKRPQGEDFMRLPMLFSLDNPNPKCKGKGAHYSASVNI 1025

Qy 1039 TSDGQ-VLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPFAFEPFYT 1097

Db 1026 LGNGSGVGATYFMTYHTVLQASADFIDAMQKARLIASNIT---RTMGLEASSYRVFPYSV 1082

Qy 1098 TNVFEYQYLITLPEGLFMLSCLIVPTFAVSCILILGLDLSRGLNLLSIVMLIVDTYVGFWA 1157

Db 1083 FVYFYEQLTVDDTIFNLGVSIGLAFITVTVLMGCELWATVIMCTVIAMILVNMGEVMW 1142

Qy 1158 LWDISNVAIVNLINLSAVGMSVEFVSHITRSFAISTKPTWLERAKKATISMGSAVAGVA 1217

Db 1143 LMGISLNAVSLVNLVMSCGSIVSEFCFSHTITRAFTLTSGSRVDRAEALAHMGSSVFSGIT 1202

Qy 1218 MTNLPGLVLGLAKAQLOIFFRFLNLLITLGLLHGLVFLPVLVSVGPDVNPALAL 1275

Db 1203 LTKFGGIVVLAFAKSOIQIFTFRMYLAIVLIGATHGLIFLPVLLSYIGESINKAKSL 1260

RESULT 7

Q8MKD8\_FELCA PRELIMINARY; PRT; 1276 AA.

AC Q8MKD8; DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Niemann-Pick C1.

OS Felis silvestris catus (Cat).

OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;

OC Felis.

OX NCBI\_TaxID=9685; [1]\_TaxID=9685;

RN [1] NUCLEOTIDE SEQUENCE.

RP MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;

RA Somers K.L., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A., Thrall M.A.;

RT "Mutation analysis of feline Niemann-Pick C1 disease.";

RL Mol. Genet. Metab. 79;99-103(2003).

DR EMBL; AF503634; AM27451.1; -, mRNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008158; F:hectogog receptor activity; IEA.

DR GO; GO:0030301; P:cholesterol transport; IEA.

DR InterPro; IPR004765; NP\_C\_type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD 57M.

DR Pfam; PF02460; Patched; I.

DR TIGRFAMs; TIGR00917; 2A060601; 1.

DR PROSITE; PS50156; SSD; 1.

SQ SEQUENCE 1276 AA; 141710 MW; C2DD3496472B71EF CRC64;

Query Match 34.8%; Score 2405; DB 2; Length 1276;

Best Local Similarity 40.7%; Pred. No. 3.8e-161;

Matches 535; Conservative 232; Mismatches 461; Indels 86; Gaps 29;

Qy 7 RGWLLWALLRLRAQSEPYTHIQGYCAFDYDCGKNPGLSGMLTSLNVCUSLNTARKI 66

Db 4 RGFALGLLLLLC---PAQVLAQS--CIWYGECC---IASGD-----KRYNCRVSGPPKPL 51

Qy 67 TGDHLILLOKICRLTYGTGNTACCSAKQIVLSLEASISITKALLITRCPACSDNFVNLHCH 126

Db 52 PKDGYDLVQELCPGFFP-DNVSLLCCDVQQLTKNQLQLPLQFLSRCPSCFYNLVNLFC 110

Qy 127 NTCSPNQSILFINVTR-----VAQLGAGQLPAVAYEAFYQHSFPAEQSDSCSRVRVPAAT 182

Db 111 LTCSPRQSQFLNVTATEDYVDPVTNQTKNVKELEYIGESFANAMYNACRDVEAPSSND 170

Qy 183 LAVGTWCVGYGSALCNAQRWLNFGQDTGNGLAPLDT--PHLEPGQAVSGIQPLNEGV 240

Db 171 KALGLLCGKDAEA-CNATNWIETFNKNDGQAPFTTIPFSDLPT-----HGMPEMNNAT 224

Qy 241 ARNESQGDVATCSQDCQAASC-----PAIARPOALDSTFYLGQMPGSLVLIIL 291

Db 225 KGDESVDVETGFCSCQDCSIVCGPKPQPPPPVPMRILGLDAMYIMMTYMAFLIVFP 284

Qy 292 CSVFA-----VVTILLVGRVA---PARDKSKMVDPKKGTSLSKLSFSTHTLLGQ 339

```
Db      285 GAFFALWCYRKRYFVSEYTPIDSNIAFSVANDRGE-----ASCCDALGAAPFEGCIRR 337
Qy      340 FFGQGWVWASWPLTIIVLSVIPVVALAAGLVFTELTDTPVELMSAPNSQARSEKAFHQD 399
Db      338 LFSQWGSFCVRNPGPIIFSLAFAACSSGLVFRVTNPNVDLWSAESSQARLEKEYFDT 397
Qy      400 HFGPFFRTNQVILPAPNRSSVRYDSLLGPK-NFSGIILDLDLLELLELLELRLHLOWMS 458
Db      398 HFGPFFRTEQLIIQAPHTSAHTYQPYPSGSDVPFGPPLDLAILHQVLDTQTAIEN--ITA 455
Qy      459 PEAQNRISLODICVAPLNPNTSLYDCINSLQYFONNRLLLTANQITLMGQTSOV-- 516
Db      456 SYNNETVTLQDICVAPLSPYNK---NCTILSVLNYFQNSHSM-----DHEIGDDFFVYA 507
Qy      517 DWKDHFLYCANAPITFDGDTALALSCMADYGAPVFPFLAIGYKGYDSEAEALIMTFSL 576
Db      508 DYHTHILYCVRAPASLNDTSLHDPCLGTGPGVFPWLVGLGYDDQYNNATALVITFPV 567
Qy      577 NNPAGDPRLAQAOKWEAELEENRAFORRMAGMFQVTTFAERLSLEDEINKTAAEDLPF 636
Db      568 NNYNDYTERLOKAHWEKEFINVKYKN---PNLTISFTTTERSIEDELANRESNGDIFTV 624
Qy      637 ATSYIVILYLSLALGSSSRVMDSKATLGLGGVAVLGVAMAAWGPESYLGIRSSL 696
Db      625 IISYAIMFLYISIALGHKICSRLLVDKISLGIAGILIVLSSVACSLGIPSYGILPLTL 684
Qy      697 VILQWVPLVLSVGADNIFIVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEAI 756
Db      685 IVIEVILVLAVGVDNIFILVQYQDERLHGETLQOQLGVLGEVAPSMFLSSFEAV 744
Qy      757 CFFLGALTPMPAVTFTALTGLAVILDFLLQMSAFVALLSDSKRQERASRLDVCCKVKP 816
Db      745 AFFLGAISKMPAVTFTSLFAGWAVLIDPLQITCFVSLGLDILKROEKRLDVLCCVRGS 804
Qy      817 ELPPGQ--GEGLLGFCQKAYAPFLHWTGCVLLFLALFGVLSYMSCHISVGLDQEL 875
Db      805 EDGTSVQASECLFRLFKHSYSPLLDKWMRPVIAIPVGLVSFVAVLNKVEIQLDQSL 864
Qy      876 ALPKDSYLLDYFLFLNRYFYGAPVYFVTTLGYNFSSEAGNAICSSAGCNFFSTQKIQ 935
Db      865 SMPDSSYVMDYFKSL-KYLHAGPPVYFVLEBGHDYTSLKQNMVCGMGNCNDSLVOQIF 923
Qy      936 YATFPPEQSYLAIPASSWDDPIDWLP-SSCRLYISGPKDKFCPSYVNSLCLNKM 994
Db      924 NAAQLDSYTRIGFAPSSIDDYFDWVKPQSSCCRYSN---TDRFCNASVVDPAICR-CR 979
Qy      995 SITM-GSVRPSVEQPHKYPFLNDRPNIKCPKGLAAYTSVN-LTSDGOVLASRFWAY 1052
Db      980 PLTQEGKQRPQGGDFMRFLPFLSDNPNPKCGKGHAAYSAVNLLGNDTGVGATYFTY 1039
Qy      1053 HKPLKNSQDYTEALRAARELANITADLRKVPGTDP-A-FEVPYPTITNVFYEQYLTLPE 1111
Db      1040 HTVLQTSADFTDAMEKANLIASNIT---KTMGLEGSNRYFVFSVFVFEQYLTIID 1095
Qy      1112 GLFMLSCLVPTFAVSCLLGLDLBSGLNLSIWMILVDTVGFALWDISYNAVSLNL 1171
Db      1096 TIFNLVSLSGAIFLTVTLLGCDLWSAVIMCITAMILVNNMFGVMWMLGSLNVAVSLNL 1155
Qy      1172 VSAGCMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGLVLGLAK 1231
Db      1156 VMSCGISVEFCSHITRATVSMKGSRAQAREALAHMGSSVFSGLITLKFQGIIVLAFAK 1215
Qy      1232 AQLIQIFFRNLMLITLLGLHLVFLPVLISYVGPDPVNPALAL---EQKRAE 1282
Db      1216 SQIFQIFFRYLANVLLGATHGLIFLPLVLLSYIGFSINKAKSLATQBOYKGE 1269
```

RESULT 8  
Q9GK52 CANFA  
ID Q9GK52 CANFA PRELIMINARY;  
AC Q9GK52  
DT 01-MAR-2001 (TREMREL. 16, Created)

```
DT      01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE      Niemann-Pick type C1 disease protein.  
GN      Name=NPCL1;  
OS      Canis familiaris (Dog).  
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC      Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
OC      Canis.  
OX      NCBI_TaxID=9615;  
RN      [1]  
RP      NUCLEOTIDE SEQUENCE.  
RA      Murakawa M., Freeman M.W.;  
RL      Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR      EMBL; AF315034; AAC40873.1; -, mRNA.  
DR      Ensembl; ENSCAFG00000018183; Canis familiaris.  
DR      GO; GO:0016021; C:integral to membrane; IEA.  
DR      GO; GO:0005764; C:lysosome; IEA.  
DR      GO; GO:0016020; C:membrane; IEA.  
DR      GO; GO:0008158; P:hedhog receptor activity; IEA.  
DR      GO; GO:0030301; P:cholesterol transport; IEA.  
DR      InterPro; IPR004765; NP_C type.  
DR      InterPro; IPR003392; Patched.  
DR      InterPro; IPR000731; SSD 5TM.  
DR      TIGRfams; TIGR00917; 2A060601; 1.  
DR      PROSITE; PS0156; SSD; 1.  
SQ      SEQUENCE 1276 AA; 141695 MW; 07D1327B51BD5681 CRC64;  
  
Query Match 34.8%; Score 2404.5; DB 2; Length 1276;  
Best Local Similarity 40.4%; Pred. No. 4.2e+161;  
Matches 524; Conservative 233; Mismatches 434; Indels 105; Gaps 28;  
  
Qy      33 CAFYDEGKNPELSGLMTLSNVCSLNTNPARKITGHLILLQKICRLTYGPNYQACS 92  
Db      25 CWVYGECC---IASG-----KRYNCQYSGPPKPLKDGIDLMQSLCPLFP-DNVSVCCD 76  
Qy      93 AKQVLSLEASLITKALLTRCPACSDNFVNLHCHTCSPNQSLFINVTR-----VAQLGAG 148  
Db      77 VQQRUTKDSQLPLQFLSRCPSCFYNNMLFCELTCSQRSQFLNTEVEDYDPTNQ 136  
Qy      149 QLPVAVAYEAFYQHSFAEQSYDSCSRVVRPAAATLAVGTWGVYGSALCNAQRWLNFGD 208  
Db      137 TKTNVKELQYVGVGSFANAMYNACRDVEAPSSNDKALGLLCGKEAE--CNATNIIEYMFN 195  
Qy      209 TGNGLAPL-----DITFHLLPEQAVGSGIQPLNEGVARCNESQGDVATCSQCCAA 261  
Db      196 KDNGQAPFTIIPISDLPAH-----GMKPMNNAATKGDPEFDEVTAPCSQCQCSV 245  
Qy      262 SC-----PATARQALDSTFYLQMPGSLVLIILCSVFAVVVITLLYGFVRVAPARD 312  
Db      246 VCGPKPQPPAPAPWRILGLDAMVIMWITYMAFLMFFGGAFFAVWCY----- 293  
Qy      313 KSKMVDPKGTSLSDKLSFSFHT-----LLGQFFQG-----WGTWVASWPLT 354  
Db      294 -RKRYFVSEVTPIDSNIAFSVWAGDTGCEASCCDGAFAFEGCLRLFTQWGSFCIRNPGC 352  
Qy      355 ILVLSVIPVVALAAGLVFTELTDTPVELMSAPNSQARSEKAFDHOHGPFRTNQVILTA 414  
Db      353 IIFSLAFIAACSSGLVFSRVTTNPNVDLWSAPGSOARLEKEYFDFAHGFPFFRTEQLIIQA 412  
Qy      415 PNRSRYVDSLILGPK-NFSGIILDLDLLELLELLELQERLHRLHQLVMSPEAQRNISLQDICYA 473  
Db      413 PHTSVHYQPYPSGSDVPFGPPLDIGILHQVLDTQTAIEN--ITATYNNETVTLQDICVA 470  
Qy      474 PLNPDTSLYDCCINSLLOYPQNRTLLLTANOTLMGQTSQV--DWKDHFLYCANAPLT 531  
Db      471 PLSYPNK---NCTIMSVLNYFQNSHSM-----DHKIGDDFVYVADYHTHLLYCVRAPAS 522  
Qy      532 FKDGDTALALSCMADYGAPVFPFLAIGYKGYDSEAEALIMTESLNINYPAGDPRLAQAOK 591  
Db      523 LNDTSLHLDPCLTGTFGGVFPWLVGLGYDDQYNNATALVITFPVNTYNDTEKLQRAQA 582  
Qy      592 WEEAFLEEMRAFQRRMAGMFQVTTFAERLSLEDEINRTTAEDLPATSIYVIFLYISIAL 651
```

Db 583 WEKEFINVKYEN---PNTISFTTRESNGDVFVTLISYAVMELYISIAL 639  
Qy 652 GSYSSWGRVVDKATLGLGVAVVLGAVMAAGFFSYLGRSSRLVLQVVPFLVLSVGA 711  
Db 640 GHKSCSRFLVDSKISGIAGILVLSVMCSGIFSYFGIPLTLVIEVPLVLAVG 699  
Qy 712 DNIPIFVLEYQRIIPRRGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRT 771  
Db 700 DNIPIFVLEYQRIIPRRGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRT 759  
Qy 772 FALTSLAVILDELQMSAFVALLSLDSKQESRLDVCCVXPQELPPPGQG-----EGL 827  
Db 760 FSLFAGVAVILDELQITCFVSLGLDLKQEKRLDLVCLLTGSE---GGTGIQASESC 816  
Qy 828 LLGFFQKAYAPFLLHMTITRGVLLFLALFGVSLYSMCHTSVGLDQELALPKDSYLLDYF 887  
Db 817 LFRFFKNSYSPFLKQWMDRPIVAVFGIILSFIANLKVIEGLDQSLSPDDSYMDYF 876  
Qy 888 LFLNRYEVGAPYFVFTTLGYNFSSEAGMNAICSSAGCNFNFSFTQKIQYATBPEOSYLA 947  
Db 877 KSL-KYLHAGFPVYFVLEEGHDYTSLEGQNMVCGMGCNDSLVQQIFSAQLDNYTHIG 935  
Qy 948 IPASSWDDPIDMLTP-SSCRLYISGPNKDKFCSTPSTVNSLNCCLKMCSITM-GSVRPSV 1005  
Db 936 FAPSSWIDDYFDWVKPQSSCCRYNS---TDQFCNASVVDPAVR-CRPLTQEGKRRPQG 991  
Qy 1006 EQPHKYLWFLNDRPNKCPKGLAAYSTSVNLT-SDQVQLASRFWAYHKPLANSODYTE 1064  
Db 992 EDFWFLPWLNSPNPKCKGKGAAYGSANLVGNDTSVGATFYHTVLTQTSADFTD 1051  
Qy 1065 ALBAARELANITADLRKRVGTDPAFEVFPYTTITNVEQYLTLPEGLFMLSLCLVPTF 1124  
Db 1052 AMRKAILIASNITKTM-SLKGSH--YRPPSYFVYFYEQYLTLIIDTIFNLVSLGAIF 1108  
Qy 1125 AVSCLLGLDLRSLNLISIVMLVDTVGFMALWDISYNAVSLINLVAAGVMSVFVSH 1184  
Db 1109 LVTVLVLGCELWSAVTMCVTIAMILVNMFGVWMLWGISLNAVSLVNLVMSCGISVDFCSH 1168  
Qy 1185 ITRSPAISTKPTWLERAKETATSMGSAVAGVAMTLPGLVLGLAKAQLIQIFPRLNL 1244  
Db 1169 ITRAFVTSAGSRVERAEALSHGSSVFGITLTFGGLIVLAFAKSLQIFQVFTGMYL 1228  
Qy 1245 LITLGLGLHGLVLPVLSVGVDPVNPALAL-EQKR 1279  
Db 1229 AMVLGATHGLIFLPVLLSVIGPSINKAKSLAQER 1264

## RESULT 9

NPCL HUMAN  
ID NPCL HUMAN STANDARD; PRT; 1278 AA.  
AC O15118; Q9P130;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 10-MAY-2005 (Rel. 47, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Niemann-Pick C1 protein precursor.  
GN Name=NPCL1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE, VARIANT ILE-642, AND VARIANTS NPCL1  
RX MEDLINE=97362323; PubMed=9211849; DOI=10.1126/science.277.5323.228;  
RA Carstea E.D., Morris J.A., Coleman K.G., Loftus S.K., Zhang D.,  
RA Cummings C., Gu J., Rosenfeld M.A., Pavan W.J., Krizman D.B.,  
RA Nagle J., Polymeropoulos M.H., Sturley S.L., Ioannou Y.A.,  
RA Higgins M.E., Comly M., Cooney A., Brown A., Kaneski C.R.,  
RA Blanchette-Mackie E.J., Dwyer N.K., Neufeld E.B., Chang T.-Y.,  
RA Liscum L., Strauss J.F. III, Ohno K., Zeigler M., Carmi R., Sokol J.,  
RA Markie D., O'Neill R.R., van Diggelen O.P., Ellender M.,  
RA Patterson M.C., Brady R.O., Vanier M.T., Pentchev P.G., Tagle D.A.;

RT "Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis." Science 277:228-231(1997).  
RL [2]  
RN NUCLEOTIDE SEQUENCE, AND VARIANTS.  
RX MEDLINE=99355599; PubMed=10425213; DOI=10.1006/bbrc.1999.1070;  
RA Morris J.A., Zhang D., Coleman K.G., Nagle J., Pentchev P.G., Carstea E.D.;  
RA "The genomic organization and polymorphism analysis of the human Niemann-Pick C1 gene." Biochem. Biophys. Res. Commun. 261:493-498(1999).  
RL [3]  
RN NUCLEOTIDE SEQUENCE, AND VARIANT ILE-642.  
RX MEDLINE=21623216; PubMed=11754101; DOI=10.1002/humu.10016;  
RA Bauer P., Knoblich R., Bauer C., Finckh U., Hufen A., Kropp J., Braun S., Kustermann-Kuhn B., Schmidt D., Harzer K., Rolfs A.;  
RT "NPC1: complete genomic sequence, mutation analysis, and characterization of haplotypes." Hum. Mutat. 19:30-38(2002).  
RL [4]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT ILE-642.  
RC TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G., Strausberg R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL [5]  
RN CHARACTERIZATION.  
RX MEDLINE=99128318; PubMed=9927649; DOI=10.1073/pnas.96.3.805;  
RA Watari H., Blanchette-Mackie E.J., Dwyer N.K., Glick J.M., Patel S., Neufeld E.B., Brady R.O., Pentchev P.G., Strauss J.F. III;  
RT "Niemann-Pick C1 protein: obligatory roles for N-terminal domains and lysosomal targeting in cholesterol mobilization." Proc. Natl. Acad. Sci. U.S.A. 96:805-810(1999).  
RL [6]  
RN VARIANT NPD TRP-992.  
RX MEDLINE=98299797; PubMed=9634529;  
RA Greer W.L., Riddell D.C., Gillan T.L., Girouard G.S., Sparrow S.M., Byers D.M., Dobson M.J., Neumann P.E.;  
RT "The Nova Scotia (type D) form of Niemann-Pick disease is caused by a G3097->T transversion in NPC1." Am. J. Hum. Genet. 63:52-54(1998).  
RL [7]  
RN VARIANTS NPCL1 GLN-934; LEU-940; ASN-948; TRP-992; ALA-1007; THR-1061 AND VAL-1213.  
RX MEDLINE=99452586; PubMed=10521290;  
RA Greer W.L., Dobson M.J., Girouard G.S., Byers D.M., Riddell D.C., Neumann P.E.;  
RT "Mutations in NPC1 highlight a conserved NPC1-specific cysteine-rich domain." J. Hum. Genet. 65:1252-1260(1999).  
RL [8]  
RN VARIANT NPCL1 THR-1061.  
RX MEDLINE=99452593; PubMed=10521297;  
RA Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A., Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;

"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant allele in patients of Western European descent and correlates with a classic juvenile phenotype.";  
Am. J. Hum. Genet. 65:1321-1329(1999).  
[9]  
RN VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GLN-1266.  
RX MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004390051057; Zhang H., Yamamoto T., Namba E., Ninomiya H., Higaki K., Taniguchi M., Zhang H., Akaboshi S., Watanabe Y., Takeshima T., Inui K., Okada S., Tanaka A., Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G., Ohno K.;  
RA "NPC1 gene mutations in Japanese patients with Niemann-Pick disease type C.";  
RT Hum. Genet. 105:10-16(1999).  
RL [10]  
RN VARIANTS NPC1 GLN-958 AND ALA-1007.  
RX MEDLINE=21313111; PubMed=11349231;  
RA Sun X., Marks D.L., Park W.D., Wheatley C.L., Puri V., O'Brien J.P., Kraft D.L., Lundquist P.A., Patterson M.C., Pagano R.B., Snow K.;  
RA "Niemann-Pick C variant detection by altered sphingolipid trafficking and correlation with mutations within a specific domain of NPC1.";  
RT Am. J. Hum. Genet. 68:1361-1372(2001).  
RL [11]  
RN VARIANTS NPC1 ALA-378; MET-950; ARG-992 AND THR-1061.  
RX MEDLINE=21313105; PubMed=11333381;  
RA Millat G., Marçais C., Tomasetto C., Chikh K., Fensom A.H., Harzer K., Wenger D.A., Ohno K., Vanier M.T.;  
RA "Niemann-Pick C1 disease: correlations between NPC1 mutations, levels of NPC1 protein, and phenotypes emphasize the functional significance of the putative sterol-sensing domain and of the cysteine-rich luminal loop.";  
RT Am. J. Hum. Genet. 68:1373-1385(2001).  
RL [12]  
RN VARIANTS NPC1 TYR-177; CYS-978 AND VAL-1035.  
RX MEDLINE=21372069; PubMed=11479732; DOI=10.1007/s004390100531;  
RA Ribeiro I., Marcao A., Amaral O., Sa Miranda M.C., Vanier M.T., Millat G.;  
RA "Niemann-Pick type C disease: NPC1 mutations associated with severe and mild cellular cholesterol trafficking alterations.";  
RT Hum. Genet. 109:24-32(2001).  
RL CC -!- FUNCTION: Involved in the intracellular trafficking of cholesterol. May play a role in vesicular trafficking in glia, a process that may be crucial for maintaining the structural and functional integrity of nerve terminals.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late endosomes and lysosomes.  
CC -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain containing a di-leucine motif necessary for lysosomal targeting.  
CC -!- PTM: Glycosylated  
CC -!- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease type C1 (NPC1) [MIM:257220]. NPC1 is an autosomal recessive lipid storage disorder, which affects particularly the brain, liver and spleen, and which is characterized by lysosomal accumulation of low density lipoprotein derived cholesterol. Clinical features include variable hepatosplenomegaly and severe progressive neurological dysfunction such as ataxia, dystonia and dementia. The age of onset can vary from infancy to late adulthood.  
CC -!- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease type D (NPD) [MIM:257220]; also known as Niemann-Pick disease without sphingomyelinase deficiency, or Nova Scotian type. Because of evidence from biochemical changes, lack of complementation, and linkage mapping to the same chromosome site, NPD and NPC1 are considered to be allelic disorders.  
CC -!- SIMILARITY: Belongs to the patched family.  
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; AF002020; AAB63982.1; -; mRNA.  
DR EMBL; AF157379; AAD48006.1; -; Genomic DNA.  
DR EMBL; AF157365; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157366; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157367; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157368; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157369; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157370; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157371; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157372; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157373; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157374; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157375; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157376; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157377; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF157378; AAD48006.1; JOINED; Genomic DNA.  
DR EMBL; AF338230; AAK25791.1; -; Genomic DNA.  
DR EMBL; AF123046; AAT28875.1; -; Genomic DNA.  
DR EMBL; BC063302; AAH63302.1; JOINED; Genomic DNA.  
DR Ensembl; ENSG00000141458; Homo sapiens.  
DR HGNC; HGNC:7897; NPC1.  
DR MIM; 607623; -.  
DR MIM; 257220; -.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0005764; C:lysosome; TAS.  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0005478; F:intracellular transporter activity; TAS.  
DR GO; GO:0015248; F:sterol transporter activity; TAS.  
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
DR InterPro; IPR004765; NP\_C\_type.  
  
Query Match 34.8%; Score 2402.5; DB 1; Length 1278;  
Best Local Similarity 39.8%; Pred. No. 5.8e-161;  
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;  
  
Qy 7 RGWLLWALLLRLAQSEPTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59  
Db 4 RGLALGALLLLCPAQVFSQ-----SCVYGEGLAYGDKRYNCEYSG----- 46  
Qy 60 NTPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASISITKALLTRCPACSDN 119  
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKDLNLQPLQFLSRCPSCFYN 103  
Qy 120 FVNLIHCNTCSPNQSLEPINVTR-----VAQLGAGOLPAVAYEAFYQHSFABQSDVCSRV 175  
Db 104 LLNLFCBLTCSPRQSQFLNVTATEDYVDPTNQTNTNKLQYVVGQSFANAMYNACRDV 163  
Qy 176 RVPAAATLAVGTWCVGVYGSALCNAQRWLNFGDGTGNGLAPLDIT-----FHLLPQQA 228  
Db 164 EAPSSNDKALGLLCCGDADA-CNATNWIETFNKNDQAPFTITPVSDFPVH----- 215  
Qy 229 VGGIQLPANGVARGCNESQGGDDVATCSQDCAASCAPAIARPO-----ALDSTFYL 279  
Db 216 ---GMEPMNATKGCDSVDEVTAPCSQDCSICVCGPKPQPPPPAPPTILGLDAMYVIM 272  
Qy 280 QMPGSLVLIILLCISVFAVITILL-----VGRVAPARDKSKWDPKKGTSLS 326  
Db 273 WITTMALFFVFGFAFAVCYRKYRYFSEVTPIDSNIAFSV-NASDKGE-----ASCC 324  
Qy 327 DKLSFSTHTLIGQFFQCGWGWASWPLTILVLSVIVVLAAGLVFTETLTDPVELWSAP 386  
Db 325 DPTVSAAEFGCLRLFTWGSFCVRNPGCVIFFSFLVFTACSSGLVFRVTNPNVDLSAP 384  
Qy 387 NSQARSEKAFHDQHFPGFFRTNQVILTPAPNRSYRYDSLLGPK-NFSGILDLDLLELL 445  
Db 385 SSOARLEKEYFDQHFSGFFRTEQLIIRAPLTDKHIYQYPSPGADVPFGPPLDIQLHQVL 444  
Qy 446 ELQERLRHLQVSPQAEARNLSLODICVAPLNPNSTLSYDCINSLLQYFQNNRTLLTLTA 505  
Db 445 DLQAIEN--ITASYDNETVTQLQICLAPLSPYNT---NCTILSVLNYFNQSHVLDHKK 499  
Qy 506 NQTLMGQTSQVDWKDHFPLYCANAPLTPKDGTLALSCWADYGAPVFVFPFLAIGGYKGDYS 565

Db 500 GDDFF---VVADYHTELYCYRAPASLNDTSLHDPCLGTFGGVPFWLGGVDDQVYN 556  
Qy 566 EAEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFOVTFTAERSLDEI 625  
Db 557 NATALVITFPVNNYNDTEKLQRAQAEKEFINFKYKN---PNLTSTFAERSIDEL 613  
Qy 626 NRTTAEDLPATSYIVIFLYISIALGYSYSWSRWVDSKATLGLGGVAVVLGAVMAAMG 685  
Db 614 NRESDSDFVTWISYAIMFLYISIALGHMKSRLVDSKVSGLGIAGILVLSVACSLG 673  
Qy 686 FFSVLGRSSIVILQVVPFLVSLVACADNIFVLEYQRLPRRPGEPREHIGRALGRVAP 745  
Db 674 VFSYIGLPLTLIVIEPFLVAVGVNDIIFILVQYQDERLOGETDQOGLGRVLGVAP 733  
Qy 746 SMLLCSLSEAICFFPLGALTTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKROEAS 805  
Db 734 SMFLSSSEVAVFPLGALSVMVPAVHTFSLFAGLAVFIDFLQITCFVSLGLDQIKRQKN 793  
Qy 806 RLDVCCCKPOELPPGQ-GEGLLGGFPQKAYAPFLHWTIRGVVLLFLALFVGLSYLM 864  
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMPVIAIFVGVLSFSTAVL 853  
Qy 865 CHLSVGLDQSLPKDSYLLDYFLFNRYFEVGAOPYVFTTLGVNFSSEAGMAICSSAG 924  
Db 854 NKVDIGLQSLMPDDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSRRQNMVCGMG 913  
Qy 925 CANNFSPTQKIQATPEPQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKDKFCPT 983  
Db 914 CNDLSVQOIFNAQLNDYTRIGFAPSWSWIDDFDWKVPQSSCCRV---DNITDQFCNAS 970  
Qy 984 VNSLNCNKCMWIT-MGSVRPSVEQHKYLPWFINDRPNIKCPKGLAAYSTSVN--LTS 1040  
Db 971 VDPACVR-CRPLTPEGKQRPQGDPMRPLPMLSDNPNPKCGKGHAAVSSAVNILLGH 1029  
Qy 1041 DGOVLASRFWAYKPLKNSODYTEALPAARELANITADLRKVPGTDPAPREVPYITNV 1100  
Db 1030 GTRVGATYFMTYHTLVLTQTSADFDALKKARLIASNVT-ETWINGS--AYRVPFYSVYV 1086  
Qy 1101 FYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYMLVDTVGFALND 1160  
Db 1087 FYEQYLITIDTTFNLGVSLGALFVTLVLLGELMSAVJMCATIAMLVNMFVGMWLMG 1146  
Qy 1161 ISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLERAKEATISMSAVPAGVAMTN 1220  
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKGSVEREAEALAHMSSVFSGITLTK 1206  
Qy 1221 LPGILVLGLAKAQLIQIFPRMLNLLTLGLLHGLVFLPVILSYVGPVNPA 1272  
Db 1207 FGGIVVLAFAKSQIFQIFPRMYLAMVLLGATHGLIFLPLVLLSYIGPSVYVKA 1258

## RESULT 10

Q59GR1 HUMAN  
ID Q59GR1 HUMAN PRELIMINARY; PRT; 1289 AA.  
AC Q59GR1;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Niemann-Pick disease, type CI variant (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,  
RA Ohara O., Nagase T., Kikuno F.R.;  
RT "None title."  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB209048; BAD92285.1; -; -.

FT NON TER 1 1  
SQ SEQUENCE 1289 AA; 143230 MW; 663DE27C13FD801F CRC64;  
Query Match 34.8%; Score 2402; DB 2; Length 1289;  
Best Local Similarity 39.3%; Pred. No. 6.3e-161;  
Matches 525; Conservative 239; Mismatches 452; Indels 106; Gaps 26;  
Qy 2 AEAGLRGWLWA-----LLRLAOSPEVTTTHQPGYCAPYDECG-----KNPELSG 47  
Db 3 ASAAIRGHSTARGLAGLGLLLLLCPAQVFSQ-----SCVYGECCGAYGDKRNCYSG 57  
Qy 48 SLMTLNVSCLSNTPARKITGDHLLILLOKICPLRYTGPNTQACCSAKQLVLEASLITK 107  
Db 58 -----PPKPLPKDGYDLVQELPCGFFFG-NVSLCCDVRQLQTLKDNLQLPL 102  
Qy 108 ALLTRCPACSDNFVNLCHNTCSNOSLFINVTR-----VAQLGAGQLPAVVAEAPQHS 163  
Db 103 QFUSRCPCPNLNLIFCELTCSPROSQFLNVTATEDYDPVTNQTNTNKEQLQYVVGQS 162  
Qy 164 FARQSYDSCSRVRVPAAAATILAVGTCMGVYGSALCNQORWLNFGQDTGNGLAPLDTIT--- 219  
Db 163 PANAMYNACEDVEAPSNDKALGLLCKDADA-CNATNWIEMFNKDNQGOAPFITPVFS 221  
Qy 220 -FHLLEPQAVSGIOPLNEGVARCNESQGDVATCSQCDCAASCAPAIAPQ----- 270  
Db 222 DFPV-----RGMEPMNATKGCDESVDVETAPCSQCSQCSIVCGPKPQPPPPAPWTI 273  
Qy 271 -ALDSTFYLGQMGPSLVLIILLCSEFVAVTILL-----VGRVAPARDKSKM 316  
Db 274 LGLDAMYVIMWITYMAFLVFFGAFVAVMCYRKRYFVSEYTPIDSNIAFVS-NASDKGE- 331  
Qy 317 VDPKGTSLSDKLFSFTHLLGQFQCGWGTWASWPLTILVLSVIPVVALAAGLVFTLT 376  
Db 332 -----ASCDDPVSAAFEGCLRLPFRWGSFCVRNPGCVIFFSLVITACSSGLVFRVT 385  
Qy 377 TDPVELWSPNGQARSEKAFHDQHPFPFRTNQVILTAPNRSYRYSDDLGLPK-NFSGI 435  
Db 386 TNPVDLWSPSSQARLEKEYFDQHGFFFRTEQLIIRAPLTDKHIQPYPSGADVPRGPP 445  
Qy 436 LLDLDLLELLELLEQLERLHLQVMSPEAQRNLSLODICVAPLNPDNTSLYDCCINSLOYFQ 495  
Db 446 LDIQILHQLDLQIAIEN--ITASYNDETVTTLQDICIAPLSPYNT---NCTILSVLNYFQ 500  
Qy 496 NNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTLALSCMADYAGVAPVPFLA 555  
Db 501 NHSVLDHKKGDDFF---VYADYHTELYCYRAPASLNDTSLHDPCLGTFGGVPFWLV 557  
Qy 556 IGGYKGYKDYSEAEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFOVTF 615  
Db 558 LGGYDDQNNATALVITFPVNNYNDTEKLQRAQAEKEFINFKYKN---PNLTISF 614  
Qy 616 TAERSLEDEINRTTAEDLPATSYIVIFLYISIALGYSYSWSRWVDSKATLGLGGVAV 675  
Db 615 TAERSIEDLNRSDSDVFTVVISYAIMFLYISIALGHMKSRLVDSKVSGLGIAGILI 674  
Qy 676 VLGAVMAAMGFFSYLGIRSLVILQVVPFLVSLVACADNIFVLEYQRLPRRPGEPREH 735  
Db 675 VLSVACSLGVSYGLPLTLIVIEPFLVAVGVNDIIFILVQYQDERLOGETDQO 734  
Qy 736 IGRALGRVAPSMLLCSLSEAICFFPLGALTTPMPAVRTFALTSGLAVIDLFLQMSAFVALL 795  
Db 735 LGRVLGEVAPSMFLSPSFTVAFPLGALSVMVPAVHTFSLFAGLAVFIDFLQITCFVSL 794  
Qy 796 SLDSKROEASRLDVCCCKVQELPPGQ-GEGLLGGFPQKAYAPFLHWTIRGVVLLFL 854  
Db 795 GLDIKROEKRLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMPVIAIFV 854  
Qy 855 ALFGVSLYSVMCHISVGLDQELAPKDSYLLDYFLFNRYFEVGAOPYVFTTLGVNFSSEA 914  
Db 855 GVLSTFSIIVLVKVDIGLQSLMPDDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSK 914  
Qy 915 GMAICSSAGCANNFSPTQKIQATPEPQSYLAIPASSWVDDFDWLTP--SSCCRLYISG 973

Db 915 GQNMVCGMGCCNDSLVQOIFNAAQLDNVTRIGFAPSSWIDDYFDWVKPQSSCCRV---D 971  
QY 974 PKNDKPCSTVNSLNCNKCMSIT-MGSVRSVEQFHXYKLPWFLNDRNINIKCPKGGLAAY 1032  
Db 972 NITDQFCNASVDPACVR-CRLPTGKQKQPGGDMERFLPMFLSDNPNPKCGKGHAAY 1030  
QY 1033 STSVN--LTSQGVASRFMAVHKPKNSQDYTEALRAARELAANITADLRKVPCTDPAP 1090  
Db 1031 SSVNLLGHTRGATGFTMYTHVTQTSADFDALKKARLIASNVT-ETWINGS--AY 1087  
QY 1091 EVFPYTIITNVFEQYLTITPEGLFMLSCLVPTFAVSCLLGLDLSRGLNLLSTVMTLV 1150  
Db 1088 RVFPYSVFVFEQYLTITDITFNLGSLGAIPLVTWVLLGCELWSAVIMCATIAMVLV 1147  
QY 1151 DTVGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMG 1210  
Db 1148 NMFGVMWMLGISLNAVSLNVMSCGISVEFCSHITRAFTVMKGSRYERAEALAHMGS 1207  
QY 1211 AVFAGVAMTNLPGILVLGLAKAQLIQIPFFRLMLLITLLGLHGLVPLVILSVYVGPDVN 1270  
Db 1208 SVFSGITHTKGGIVVLAFAKSQIFQIFFRMYLANVLLGATHGLFIPVLSYIGPSVN 1267  
QY 1271 PA 1272  
Db 1268 KA 1269  
RESULT 11  
ID Q7TMD4 MOUSE PRELIMINARY; PRT; 1277 AA.  
AC Q7TMD4.  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 10-MAY-2005 (TREMELrel. 30, Last annotation update)  
DE Nienann Pick type C1.  
GN Name=Npcl;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahay J., Helton E., Kettenan M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC052437.1; -; mRNA.  
DR EMBL; BC054539; AAH54539.1; -; mRNA.  
DR GO; GO:0006897; P: endocytosis; IMP.  
DR InterPro; IPR004765; NP\_C type.  
DR InterPro; IPR003392; Patched.  
DR InterPro; IPR000731; SSD 5TM.  
DR Pfam; PFO2460; Patched; 1.  
DR TIGRFAMs; TIGR00917; 2A060601; 1.  
DR PROSITE; PS0156; SSD; 1.  
SQ SEQUENCE 1277 AA; 142882 MW; 3B42230AAC8E564E CRC64;

Query Match 34.7%; Score 2400; DB 2; Length 1277;  
Best Local Similarity 40.4%; Pred. No. 8.7e-161;  
Matches 528; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY 11 LWALLLLLAQSEPEYTHIQPGYCAFYDECGNPELSGLMTLSNVSCLSNTPARKKITGDH 70  
Db 8 LGLLLLLLCPAQVESQ-----SCVWYGECC---IATGD---KRYNCKYSGPPKPLPKDG 55  
QY 71 LILQKTCPRLYTGPNTQACSAKQLVSLASISITKALLTRCPACSDNFVNHLCHNTCS 130  
Db 56 YDLVQELCPGLIFF-DNVSLLCCDIQQLQTLKSNLQLPLQFLSRCPSCFYNTLMTLFCELTCS 114  
QY 131 PNQSLFNTVTRVAQLGAGQLPA-----VVAYEAFVQHSFASQSYDCSRVRVPAATLAVG 186  
Db 115 PHQSQFLNLTATEDYFPDKTQENKTNVKELEYFVQSFANAMYNARDVEAPSNKALG 174  
QY 187 TMCVGYSGALCNAQRWLNFOQDTGNGLAPLDI-----TFHLEPGQAVGSGIQPLNEGVA 241  
Db 175 LLCGRDARA-CNATNWEYMFENKNGQAPFTIIPVPSDLSTL-----GMEPWRNATK 225  
QY 242 RCNESQGDVATCSQCQCAASC-----PATARQALDSTP-----VILQMPG 283  
Db 226 GCNESVDEVGTGPCSCQDCSICVCGPKQPPPPPPMPRIWGLDAMVIMWVTVVAELVFVFG 285  
QY 284 SLVLIILLCSPFAVVVITLLGFRVAPARDKSKMDPKGTSLSDKLSFSTHTLLGQFPQG 343  
Db 286 ALLAVWCHRRIRYFVSEYTPIDSNJAFSVNS-----DKGEASCCDPLGAFDDCRRMPTK 341  
QY 344 WGTWVASWPLTILVSVIPVVALAAGLVFTLTDPVELMSAPNSQARSEKAFDQHPGP 403  
Db 342 WGAFCVRNPTCIIFFSLAFITVCSSGLVFVQVTTNPVELMSAPHQARLEKEYDKHFGP 401  
QY 404 FFRNQVILFAPNRSRYDSLLGPK-NFSGILDLDLLELLELQERLHLQWSPSEAQ 462  
Db 402 FFRTEQLIQAPNTSVHIYEPYPAGADVPPFGLPKKEILHQVLDLQ-----IAIESITAS 456  
QY 463 RN---ISLQDICVAPLNPDTLSLYDCCINSLLQVFNQNRLLLTLLTANQTLMGQTSQV--- 516  
Db 457 YNNTVTLQDLCVAPLSPYK---NCTIMSVNLTQNSHAVL-----DSQVGD 502  
QY 517 -----DKWDHPLYCANAPLTFKDGTLALSCMADYGAIPVFPFLAIGGYKGDYSEAEAL 571  
Db 503 FYIYADVHTHFLYCVRAPASLNDTSLHGPCJGTGFGVPVPLVLGGDDQYNNATLV 562  
QY 572 MTPSLNYPAGDPRLAQAULWEAFLEEMRAFORMMAGMFQVTTAERSLEDEINRTAE 631  
Db 563 ITFFVNNYNDTERLQRAWEKEFISFVKNYKN---PNLTISFTAERSIEDELNRENS 619  
QY 632 DLPFATSYIVIFLYISALGSYSSWSRVMVDKATGLGAVVVLGAVMAAGFFSYLG 691  
Db 620 DVFTVILSVVMFLYISALGHQISCSRLVDSKISIGIAGILVLSVACSLGIFSYMG 679  
QY 692 IRSSLVLQVVPFLVLSVGADNIFIVLEYQRLPRRPPPEPREVHIGRALGRVASMILCS 751  
Db 680 MPLTLIVIEVIPFLVAVGVNDNIFILVQTYORDERLQEBETLDQQLGRILGEVAPTMFLSS 739  
QY 752 LSEAIICPLGNLTWPAVRTFALTSLGLAVILDLLQMSAFVALLSLQSKRGEASRLDVC 811  
Db 740 FSETSAFFFGALSSMPAVHTFSLFAGMAVLIDFLQITCFVSLGLGDIKROKHNHLDLC 799

QY 812 CVKQELPPQGG-----EGLLGPFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHI 867  
Db 800 CVRGAD--DQGGSHASESYLFRFKNYFAPLLKDWLRPIVAVFVGVLSFSAVVNKKV 856  
QY 868 SVGLDQELALPKQSYLLDYFLFLNRYFEVGAFFVFTTLGVNFSSEAGMNAICSSAGCNV 927  
Db 857 DIGLDQSLSPNDSDYIDYFKSLAQYLHSGPPVYFVLEEGYNSRRKGQNNVCGGMGCDN 916  
QY 928 FSFTQKIQYATEPPEQSYLAIPASSWVDDFIDMLTP--SSCRLYISGPNKDKFCPSTVNS 986  
Db 917 DSLVQOIFNAELDTYTRVGFAPSSWIDDYFDWVPOSSCCRLY--NVTHQFCNASVMD 973  
QY 987 LNCLEKNCMSIT-MGSVPSPVEQFHYLPWFELNDAPNICKPKGGLAAVSTSVNLTSQGV 1045  
Db 974 PTCVR-CRPLTPEGKORPOGKEFKFLPMLFSLDNPENPKCGKGHAAYGSAVNI VGD DTYI 1032  
QY 1046 -ASRFMYHKLKNSQDYETALRAARELANITADLRKVPCTOPAFEPVPTITNVVEQ 1104  
Db 1033 GATYFMYHTILKTSADYTDAMKARLIASNITETMRS-KGSD--YRVFPYSVYVFEQ 1089  
QY 1105 YLTILPGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTGVGFMALWDISYN 1164  
Db 1090 YLTIIDTIFNLSVLSGIFLVTLVGLCELSAVIMCIITAMLLVNMFGVWMLGLSLN 1149  
QY 1165 AVSLINLVAGMSVFEVSHITSFAISTPTWLERAKEATISMGSAVFAGVAMTNLPGI 1224  
Db 1150 AVSLNVLNMGSGISVEFCSHITRAFTMSTKGSVRSRAEALAHMGSSVFGGITLTKFGGI 1209  
QY 1225 LVGLAKAQLIQIIFFFELNLLITLGLLHGLVLPVLTLSVVGPDVNP 1272  
Db 1210 VVLAFAKQSFIEFYFRMYLAMVLLGATHGLIFLVLLSYGSPVNKA 1257

RESULT 12

Q9N0Q0\_FELCA PRELIMINARY; PRT; 1276 AA.  
AC Q9N0Q0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Niemann-Pick type C1 disease protein.  
GN Name=NPCI;  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;  
OC Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RA Murakawa M., Freeman M.W.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP258783; AF72187.1; -, mRNA.  
DR GO; GO:0016021; Crinintegral to membrane; IEA.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.  
DR GO; GO:0030301; P:cholesterol transport; IEA.  
DR InterPro; IPR004765; NP\_C\_type.  
DR InterPro; IPR003392; Patched.  
DR InterPro; IPR000731; SSD\_5TM.  
DR Pfam; PF02460; Patched; I.  
DR TIGRFAMs; TIGR00917; 2A060601; 1.  
DR PROSITE; PS50156; SSD; 1.  
SQ SEQUENCE 1276 AA; 141739 MW; CB9B7A0261691256 CRC64;

Query Match 34.7%; Score 2399; DB 2; Length 1276;  
Best Local Similarity 40.6%; Pred. No. 1e-160;  
Matches 534; Conservative 232; Mismatches 462; Indels 86; Gaps 29;  
QY 7 RGMWALLLRQAQSPYTIHQPGYCAFDECGKPELSGMTVNSVCSLNTNPARKI 66  
Db 4 RGPALGLLILLCL---PAQVLAQS--CIWYGECC---IASGD----KRYNCKYSGPPKPL 51

QY 67 TGDHLILLOKICPRLYTGPNTQACSAKQVLSLEASLITKALITRCPACSDNFVNLHCH 126  
Db 52 PKGYDLVQELCPGFFP-DNVSLLCCDQVQLQTLKDNLQLPLQFLSRCPFCFYNLVNLFCE 110  
QY 127 NTCSPNQSLEFINVTR-----VAQIGAGOLPAVAYEAFYQHSFAPBQSYDSCSRVRVPAAT 182  
Db 111 LTCSPRQSQFLNVTATEDYVDPVNTQTKTNVKEQLQYIGGSFANAMYNACRDVEAPSND 170  
QY 183 LAVTGMGVYGSALCNAQWLNFGQDTNGGLAPLDIT--PHLLEPGAVSGIOPLNEGV 240  
Db 171 KALGLICGDAEA-CNATNWIYFNSKDNQOAFPTITPISDLPT-----HGMPEMNAT 224  
QY 241 ARCNESGDDVATCSCODCAASC-----PAIARPOALDSTFYLGQMPGSLVLIIIL 291  
Db 225 KGCDSEVDEVTGFCSCODCSIVCGPKPQPPPPVPMRILGLDAMYVIMWTYMAFLVFPF 284  
QY 292 CSVPA-----VTVILLVGRVA---PAROKSMVDPKKGTSLSDKLSFSFTHLLGQ 339  
Db 285 GAFFALWCYRKRYFVSEYTPIDSNIAFSVNANDRGE-----ASCCDALGAAPGECIRR 337  
QY 340 FFGMGCTWASWPLTILVLSVIPVALAAGLVFETLTDPELWSPNSOARSEKAFHQ 399  
Db 338 LFSQMGSCFCVRNPGPIFFSLAFIAACSSGLVFRVTINPVDLWSAPSSQARLEKEYFDT 397  
QY 400 HFGPFRTNQVILTAPNRSRYSDLSLLGPK-NFSGILDLDLLELELELERLRLQVMS 458  
Db 398 HFGPFRTTEQLIIQAHPHTSAHTYQYPGSDVPPGPPDLAILHQLVLDLQATEN--ITA 455  
QY 459 PEQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQV-- 516  
Db 456 SYNNETVTLLQDICVAPLSPYNK---NCTILSVLNYFQNSHML-----DHEIGDDPFVYA 507  
QY 517 DWKDHLYCANAPLTFKDGTTALALSCHADYGAVPFPFLAIGGYKGDYSABALIMTFSL 576  
Db 508 DYHTHLLYCVRAPASLNDTSLHDPCLGTGFGVPFVPLVGGYDDQNNYNATALVITFPV 567  
QY 577 NNYPAGDPRLAQAKLWEEAFLEEMRAFQRRMAGMFOVFTTAERSLSEDEINRTTAEDLPF 636  
Db 568 NNYNDTERLQAHVWEKEFINVKNYKN---PNLTISFTTERSIDELELNRESNGDITV 624  
QY 637 ATSYIVIFYLSIALGYSWSRWVDSKATIGLGVAVVLGAVMAAMGFYSYLGIRSSL 696  
Db 625 IISYATMFLYISIALGHIKSCSRLVDSKISLGIAGILIVLSKACSLGIFSVMGPIPLT 684  
QY 697 VILQVVPFLVLSVGADNIEIFVLEYQRLPRRCEPREVHIGRALGVAPSMCLCSSEAI 756  
Db 685 IVIEVIPFLVAVGVNDIIFILVQYQYQDERLHGETLDDQQLGRVLGSAVPMFLSSFSEAV 744  
QY 757 CFFLGALTMPAVRTPALTSLGLAVILDFLLOMSAFVALLSLDSKROEASRLDYCCCKVPQ 816  
Db 745 APFLGALSKWPAVHTFSLFAGVAVLIDFLQITCFVSLGLGLDIKROEKRLDLVCCVRS 804  
QY 817 ELPPPPQ-GEGLLLGPFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVGLDQEL 875  
Db 805 EDGTSVQASESCLFRLFKHSYPLLLKDWMPVIAIFVGVLSFSVAVLNKVBIGLDQSL 864  
QY 876 ALPKDSYLLDYFLNRYFEVGAFFVFTTLGVNFSSEAGMNAICSSAGCNPSFTQKIQ 935  
Db 865 SMPDDSYVMDFKSL-KYLHAGPPVYFVLEEGHDYTSKQNNVCGMGNNDLSLQQIF 923  
QY 936 YATEFPEQSYLAIPASSWVDDFIDMLTP--SSCRLYISGPNKDKFCPSTVNSLNCLEKNCM 994  
Db 924 NNAQLDSYTRIGFAPSSWIDDYFDWVPOSSCCRVNS---TDRFCNASVVDPAICR-CR 979  
QY 995 SITM-GSVRPSVEQFHYLPWFELNDPNTIKCPKGLAAVSTSVN-LTSDQVLAASFMA 1052  
Db 980 PLTQEGKORPGQDFMRFLPMLFSLDNPENPKCGKGHAAYSSAVNLGNDTGVGATVFTY 1039  
QY 1053 HKPLKNSQDYETALRAARELANITADLRKVPCTDPA-FEVEPYITNTVYRQYLILPE 1111  
Db 1040 HTVLOTQDFTDAMRKANLIASNIT-----KTMWLGSGSNRYRVFPYSVYFYRQYLTIID 1095



DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Niemann-Pick C1 protein precursor.  
 GN Name=Npci;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RX [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=97362324; PubMed=9211850; DOI=10.1126/science.277.5323.232;  
 RA Loftus S.K., Morris J.A., Carstea E.D., Gu J.Z., Cummings C.,  
 RA Brown A., Ellison J., Ohno K., Rosenfeld M.A., Tagle D.A.,  
 RA Pentchev P.G., Pavan W.J.;  
 RT "Murine model of Niemann-Pick C disease: mutation in a cholesterol  
 homeostasis gene.";  
 RL Science 277:232-235 (1997).  
 RN [2]  
 RP SUBCELLULAR LOCATION.  
 RX MEDLINE=99145606; PubMed=9990080; DOI=10.1073/pnas.96.4.1657;  
 RA Patel S.C., Suresh S., Kumar U., Hu C.Y., Cooney A.,  
 RA Blanchette-Mackie E.J., Neufeld E.B., Patel R.C., Brady R.O.,  
 RA Patel Y.C., Pentchev P.G., Ong W.-Y.;  
 RT "Localization of Niemann-Pick C1 protein in astrocytes: implications  
 for neuronal degeneration in Niemann-Pick type C disease.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1657-1662 (1999).  
 CC -!- FUNCTION: Involved in the intracellular trafficking of  
 cholesterol. May play a role in vesicular trafficking in glia, a  
 process that may be crucial for maintaining the structural and  
 functional integrity of nerve terminals.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late  
 endosomes and lysosomes.  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in perisynaptic  
 astrocytic glial processes. Also expressed in heart, spleen, lung,  
 liver, skeletal muscle, kidney, testis.  
 CC -!- INDUCTION: Activated by the drugs progesterone and U-18666A which  
 block cholesterol transport out of lysosomes and by the  
 lysosomotropic agent NH4Cl.  
 CC -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain  
 containing a di-leucine motif necessary for lysosomal targeting  
 are critical for mobilization of cholesterol from lysosomes.  
 CC -!- SIMILARITY: Belongs to the patched family.  
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC -----  
 DR EMBL; AF003348; AAB63372.1; -; mRNA.  
 DR EMBL; AF003349; AAB63373.1; -; Genomic DNA.  
 DR FTR; J30188; J30188.  
 DR Ensembl; ENSMUSG0000024413; Mus musculus.  
 DR GO; GO:006897; P: endocytosis; IMP.  
 DR InterPro; IPR004765; NP\_C type.  
 DR InterPro; IPR003392; Patched.  
 DR InterPro; IPR000731; SSD\_5TM.  
 DR Pfam; PF02460; Patched; 1.  
 DR TIGRFAMs; TIGR00917; 2A060601; 1.  
 DR PROSITE; PS0156; SSD; 1.  
 KW Glycoprotein; Lysosome; Signal; Transmembrane.  
 FT SIGNAL 1 23 Potential.  
 FT CHAIN 24 1278 Potential.  
 FT TRANSMEM 271 291 Potential.  
 FT TRANSMEM 352 372 Potential.  
 FT TRANSMEM 623 643 Potential.  
 FT TRANSMEM 655 675 Potential.  
 FT TRANSMEM 685 705 Potential.  
 FT TRANSMEM 761 781 Potential.  
 FT TRANSMEM 834 854 Potential.  
 FT TRANSMEM 1099 1119 Potential.

FT	TRANSMEM	1125	1145	Potential.
FT	TRANSMEM	1196	1216	Potential.
FT	TRANSMEM	1228	1248	Potential.
FT	DOMAIN	621	786	SSD.
FT	MOTIF	1275	1278	Di-leucine motif.
FT	COMPBIAS	250	260	Poly-Pro.
FT	CARBOHYD	71	71	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	123	138	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	186	186	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	223	223	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	229	229	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	415	415	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	460	460	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	525	525	N-linked (GlcNAc. . .) (Potential).
FT	CONFLICT	474	474	S -> F (in Ref. 1; AAB63373).
FT	CONFLICT	479	479	N -> D (in Ref. 1; AAB63373).
SQ	SEQUENCE	1278 AA; 142889 MW; 43C71CE47D283674 CRC64;		

Query Match 34.5%; Score 2385; DB 1; Length 1278;  
 Best Local Similarity 40.2%; Pred. No. 1e-159;  
 Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY	14	LLRLAQSEPTTIHOPGYCAFYDECGKNPELGSMLTSLNSVCLSNTPARKITGDHLIL	73
DB	12	LLLLCPAQVFSQ-----SCVWYGECC-----IATGD-----KRYNCKYSGPKPLPKDGYDL	59
QY	74	LQIKCPRLYTGPNTOCCSAKQVLSLEASITKALLTRCPACSDNFVNHLCHNTTCSNQ	133
DB	60	VOELCPGLFF-DNVSLCCDIQQLTKLSNLQLPLQLSLRCPSCFYNLMTLFCELTCSPHQ	118
QY	134	SLFINTVRAQLGAGQLPA-----VVAEAFYQHSFAEQSYDCSRVVRPAANTLAVTMC	189
DB	119	SOFLNATATEDYDPKTPENKTNVKELYVGOSFANAMYNACRDVEAPSSNEKALGLIC	178
QY	190	GVYGSALCNAQRWLNFGDTGNGLAFLDI-----TFHLLPQGAQVSGIGIPLNEGVARCN	244
DB	179	GRDARA-CNATNWIEYFNKONGQAPFTIIPVPSDSLIL-----GMEPMENATKGCN	229
QY	245	ESGGDDVATCSCQDCAASC-----PAIARPAQALDSTF-----YLGOMPQSLV	286
DB	230	ESVDEVTPGSCQDCSIVGCPKPPPPPPMPWRIGLDAMVYIMVYVAFVFFGALL	289
QY	287	LIILLCSPANVTILLVGRFVAPARDKSKVDPKGTSLSDKLSFSTHTLLGQFFQWGT	346
DB	290	AVWCHRRYFVSEYTPIDNSIAFSVNSS-----DKGEASCCDPLGAADFDCLERMFTKWA	345
QY	347	WVASWPLTILVLSVIVWALAAGLVFTLTDPVELWSAPNSQARKEKAFHDOHFGPFR	406
DB	346	FCVRNPTCIIFSLAFITVCSGLVFVQVTTNPVELWSAPHQARLEKEYFDKGFDPFR	405
QY	407	TNQVILTAPNRSRYDLSLLGPK-NFSGILDLDLLELELELERLHRHQVWSPQAQRN-	464
DB	406	TEQLITQAPNTSVHIEYPAGADVFPGLNKEILHQVNLQ-----IAIESITASYN	460
QY	465	--ISLQDICVAPLNPDNTSLYDCCINSILQYQNNRTLLLTANQTLMGTSQV-----	516
DB	461	ETVTLQDICVAPLSPYNK---NCTMSVNLVFNQNSHAVL-----DSQVGDDFYI	506
QY	517	--DWKDFLYCANAPLTFKDGATALASCMDADYGAFFPFPFLAIGYKGYKDYSEALIMTF	574
DB	507	YADYTHFLYCYRAPASLNDTSLHGPCGLGTGGVPFWLVGGYDDQYNNATYLVITP	566
QY	575	SLNNYPAGDPRLAQAKLWEEAFLEENRAFORRMAGMFQVTTFAERSLEDEINRTAEDLP	634
DB	567	PVNNYNDTERLQRAWAKEFISFVKNYK---PNLTISPTAERSIEDLNESNSDVF	623
QY	635	IFATSYIVFLYISALGYSYSSRWVWDKATLGLGVAVVLGVAAMGVPFSYLGIRS	694
DB	624	TVIISVWVFLYISALGHIQCSRLVDSKISLGTAGILIVLSSVACSLGIFSYNGMPL	693
QY	695	SLVILQVWVPLVLSVGCADNIFIFVLEYQRLPRRPGEPREHVHIGRALGRVAPSMILCLSE	754



Db	994	DFMRFLPMFLSDNPNPKCGKHGAAYSAANILDNTRVGATYFMTYHTVLQTSADFIDA	1053
Qy	1066	LRAARELAANITADLRKVPGTDPAFEFVFPYITITNVFYEQYLTILPEGLFMLSCLVPTPA	1125
Db	1054	MEKARLIASNITRTMNOQGD--HRVFYSVVFVFEQYLTWIDDTIFNLSVSLGAIFL	1110
Qy	1126	VSCLLGLDLRSGLLNLLSIWMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHI	1185
Db	1111	VAVVLLGCELWSAVIMCATIAMILVNMFGVWMLWGISNAVSLVNLVMSCGISVEFCSHI	1170
Qy	1186	TRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAOLIOIFFPRLLLL	1245
Db	1171	TRAPTSTKGRVERAEEALSHMGSSVFSGITLTGFGIITVLAPAKSQIFQIFYFRMYLA	1230
Qy	1246	ITLLGLHLGLVFLFVLISVVGPDVNPALAL	1275
Db	1231	MVLGATHGLIFLPVLLSYIGPSINKAKSL	1260

Search completed: April 7, 2006, 19:21:05  
Job time : 274 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.7  
Copyright (C) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:21:19 ; Search time 52 Seconds  
(without alignments)  
2117.769 Million cell updates/sec

Title: US-10-736-769-4  
Perfect score: 6909  
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNGRQF 1332

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5-COMB.pep:\*

2: /cgn2\_6/ptodata/1/iaa/6-COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/H-COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep:\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	34.8	1318	2	US-09-949-016-10152
2	2402.5	34.8	1278	2	US-09-462-136-2
3	2385	34.5	1319	2	US-09-462-136-4
4	1329	19.2	1170	2	US-09-462-136-6
5	1078	15.6	1296	2	US-09-462-136-9
6	651.5	9.4	1447	1	US-08-540-406-19
7	651.5	9.4	1447	2	US-08-656-055-19
8	651.5	9.4	1447	2	US-08-954-668-19
9	651.5	9.4	1447	2	US-09-268-140-5
10	651.5	9.4	1447	2	US-09-918-658-19
11	651.5	9.4	1447	2	US-09-724-631-19
12	651.5	9.4	1447	2	US-08-954-701A-19
13	651.5	9.4	1447	2	US-09-807-007-6
14	651.5	9.4	1447	2	US-09-754-032-19
15	651.5	9.4	1447	2	US-08-916-140-19
16	651.5	9.4	1447	4	PCT-US95-13233-19
17	646.5	9.4	1434	1	US-08-540-406-10
18	646.5	9.4	1434	2	US-08-656-055-10
19	646.5	9.4	1434	2	US-08-954-668-10
20	646.5	9.4	1434	2	US-09-918-658-10
21	646.5	9.4	1434	2	US-09-724-631-10
22	646.5	9.4	1434	2	US-08-954-701A-10
23	646.5	9.4	1434	2	US-09-754-032-10
24	646.5	9.4	1434	2	US-08-916-140-10
25	646.5	9.4	1434	4	PCT-US95-13233-10
26	639.5	9.3	642	2	US-09-248-796A-16161
27	628	9.1	1296	2	US-08-857-636-60

28	584	8.5	1203	2	US-09-293-505-2	Sequence 2, Appli
29	584	8.5	1203	2	US-09-060-939A-2	Sequence 2, Appli
30	583.5	8.4	1182	2	US-09-293-505-7	Sequence 7, Appli
31	583.5	8.4	1182	2	US-09-060-939A-7	Sequence 7, Appli
32	581.5	8.4	531	2	US-09-248-796A-16160	Sequence 16160, A
33	580	8.4	1203	2	US-09-207-857-2	Sequence 2, Appli
34	580	8.4	1203	2	US-09-909-280A-2	Sequence 2, Appli
35	579	8.4	1146	2	US-09-807-007-1	Sequence 1, Appli
36	526.5	7.6	1311	1	US-08-540-406-4	Sequence 4, Appli
37	526.5	7.6	1311	2	US-08-656-055-4	Sequence 4, Appli
38	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
39	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
40	526.5	7.6	1311	2	US-09-724-631-4	Sequence 4, Appli
41	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
42	526.5	7.6	1311	2	US-09-754-032-4	Sequence 4, Appli
43	526.5	7.6	1311	2	US-08-916-140-4	Sequence 4, Appli
44	526.5	7.6	1311	4	PCT-US95-13233-4	Sequence 4, Appli
45	474	6.9	1286	2	US-09-268-140-3	Sequence 3, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-949-016-10152  
; Sequence 10152, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10152  
; LENGTH: 1318  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10152

Query Match 34.8%; Score 2403; DB 2; Length 1318;  
Best Local Similarity 39.7%; Pred. No. 3.9e-218;  
Matches 525; Conservative 238; Mismatches 451; Indels 110; Gaps 26;

QY	2	AEAGLRGWLWA-----LLRLAQSEPYTHIQPGYCAFYDECG-----KNPELSG 47
DB	32	ASAAALRGHSMTARGALGLLLLLCPAQVFSQ-----SCVWYGEGTAYGDKRYNCEYSG 86
QY	48	SLMTLSNVSCLSNTPARKITGDHLLILQKICPRLYTGNTQACCSAKOLVSEASLSITK 107
DB	87	-----PPKPLPKDGYDLVOELCGGFFG-NVSLCDDVRLQTLKONLQPL 131
QY	108	ALLTRCPACSNFVNLHCHNTCSPNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHS 163
DB	132	QFLSRCPSCFVNLNLFCELTCSPRSQSLNVTATEDYDVPVTKTNVKELQYVYVQS 191
QY	164	PAEQSYDSCSRVRVPAATLAVGTWCGVYGALCNAORWLNFOQDGTGNGLAPLDTIT----- 219
DB	192	FANAMYNACRDVEAPSSNDKALGLLCGKDADA-CNATNWIETMFNFKONGQAPFTITPVFS 250
QY	220	---PHLEPGQVSGIOPNEGVARCNESGDDVATCSCODCAASCAPAIARPO----- 270
DB	251	DFPVH-----GMEFPMNATKCDSEVDVETAPCSCQDCSIVCGFKPQPPPPPPAPM 300
QY	271	---ALDSTFYLQMPGSLVLIILCSVFAVVVTILL-----VGRFVAPARDKS 314





```
Qy 1162 SYNAVSLNLVSAGVMSVEFVSHITRSTFAISLTKPTWLERAKEATISMGSAVFAGVAMTNL 1221
Db 1189 SLNAVSLNLVMSGIVSEFCSHITRAFTSTKGSVRSAEELAHMGSSVFSGITUTKF 1248
Qy 1222 PGLVLGLAKAQLQIIFFRNLNLITLLGLHLGLVFLPVLISVYGVDPVNP 1272
Db 1249 GGVVLAFAKSQIFEIFFRMYLAMVLLGATHGLIFLPVLLSYIGSPVNA 1299

RESULT 4
US-09-462-136-6
; Sequence 6, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-462-136-6

Query Match 19.2%; Score 1329; DB 2; Length 1170;
Best Local Similarity 28.8%; Pred. No. 3.4e-116; Indels 182; Gaps 46;
Matches 373; Conservative 230; Mismatches 512;

Qy 10 LLWAL-----LLRLAQSEPYTHIQPGYCAFYDCGKNP-----ELSGSLMTLSNVSLSN 60
Db 3 VLMTIALVGQLMRLVQG-----TATCAMYGNCGKKSFGNELPCVPFRSFEPPVLSD 54
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTQACSAKQLVLEASLITKALLTRCPACSDNF 120
Db 55 ETSK-----LLVEVCGEEWKEVR-YACCTKDQVVALRDNLQAOPLISSCPACLKNF 105
Qy 121 VNLHCHNTCPNQSLFNIVTRVAOLGAGQLPAVVAEAFYQHSFAEQSDYSCSRVRVPA 180
Db 106 NNLFCHTCAADQGRFVNITKV-EKSKEDKDIVAEIDVFNWSSWASEFYDCKNKFSA 164
Qy 181 ATLAVGTCGVYGSALCNAQRWLNFGQDGTGN--GLAPLDTITFHLLPEQAVGSGIQPLNE 238
Db 165 N---GYAMDLIGGAKNYSQFLAKFLGDAKPMLGSPFQINIKYDLANEE--KEWQBFND 218
Qy 239 GVARNESQGDVATCQDCQCAACSPAIARQALDSTFYLGQMPG---SLVLIILCSVF 295
Db 219 EVYACDDAQ----YKACSDCQSCPHL-KP-LKDGVCVKVGLPCFSLSVLIIFYTICALF 272
Qy 296 AVVTILLVGRFVARPARDKSKMDPKGTSLSDK-----LSPSTHT-----LLGQFFQGW 345
Db 273 AFMYILCKRKNAGAMIVDDDIVESG--SLDESETNVFESFNNTNPFNGKLANLFTKVG 331
Qy 346 TWASWPLTILAVLSPVVALAAGLV-FTELTTDPVELWSPAPNSQARSEKAFHQHGF 404
Db 332 QFSVENPKLITTVFSIFVFSFIIFQVATLETDPINLWSKNSEKFEKEKYFDNFGPF 391
Qy 405 FRTNQVILTAPNRSYRSDSL--LLGPNFSGILDLDLLELLELQELRHLQWSPBAQ 462
Db 392 YRTEQIFVNETGPVLSYETLHWFFDENF-----ITEEL-----QSS 429
Qy 463 RNISLQDICYAPLNPDNLTSLDYCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDH 522
Db 430 ENIGYQDLCFRP--TEDST-----CVIESFTQYFQG-----ALPNKDSWREL 470
Qy 523 LYCANAPLTFKDGATALSCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNYPAG 582
```

```
Db 471 QECGKFP-----VNCLPFTFQOPLKTNLL---PSDDDLNNAHAFVVTLLLTNH--- 514
Qy 583 DPRLAQKLEWEEAFLEEMRAFORRMAGMFQVTTAERSLEDEINRTTAEDLPFIPATSYIV 642
Db 515 ---TQSANRMEER-LEEYLLDLKVPGL-RISFNTEISLEKELNN--NDISTVAISYLM 567
Qy 643 IFLYISIALGYSYSSWSRVVDSKATLGLGAVAVLGAVMAAMGFPSYLGIRSSIVILQV 702
Db 568 MFLYATWALRRKDGKTRLL-----LOISGLLIVLASIVCAAGFLTUFGLKSTLIIAEVI 621
Qy 703 PFLVLSVGADNIFIFVLEYQR-LPRRPGEPREVIHGRALGRVAPSMMLCSLSEASECFPLG 761
Db 622 PFLILAIGIDNIFLITHEYDRNCEQKPEYSIDQKIISAIGRMSPSILMSLLCOTGCFLIA 681
Qy 762 ALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLDSKQEQASRLDVCCKRQPELPPP 821
Db 682 AFVTMPAVHNFATYSTVSIFNGVQLQITAYVTSLSLYEKRSNYKQIT----- 728
Qy 822 GQCEGLLLGPFQKAYAPFLHLHMITRQVWLLFLALPGVSLYSYMSCHISVGLDQELALPKDS 881
Db 729 -GNEETKESFLKTFYFKMLTQ---KRLIIIIIFSAWFTSLVFLPEIOFGDLOTLAVQDS 784
Qy 882 YLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSEAGMNAICSS-AGCNNFSPFKIQYATEF 940
Db 785 YLVDYFKDYSFLNVGPPVYVMVK-NLDLTKRQKQKICGKFTTCERDLSLANVLE---QE 840
Qy 941 PEQSYLAIPASSWVDDFIDWLTSS--CCRLIYISGPNKDKFCPSTVNSLNCNKMCSITM 998
Db 841 RHRSTITEPLANWLDYFMEFLNPNQDCCRL---KKGTDEVCPSPFSPRRRC-ETCPO--Q 894
Qy 999 GSVR-----PSVEQFHKYLPWFELNDRPNIKCPKGLAAYSTSVNLTSDQGVLASRMAY 1052
Db 895 GSNVYNMSGPECKDWEYLSIWIN-APSDPCPLGGRAPYSTAL-VYNETSASVSVRTA 952
Qy 1053 HKPLKNSQDYTEALRAARELANITADLRKVPQDTPAFEVFPYITITNVFQYLTILPEG 1112
Db 953 HHPLRSQKDFQAY-----SDGVRISSSFPELDMFAYSPFYFFVQYOTLGLPT 1001
Qy 1113 LFMLSLCLVTPPAVSCLLGLDLSRLNLLLSIVMLIVDTVGPWALWDISYNAVSLNLIV 1172
Db 1002 LKLIQSAIILIFFISSVFL-QNIRSFLLALVVTMIIVDIDGALMALLGISLNAVSLNLII 1060
Qy 1173 SAVGMSVEFVSHITRSTFAI---STKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGL 1229
Db 1061 ICVGLGVFCVHIVRSETVVPSETKKDANSRVLSYSLNTIGESVIKGITLTKFVGCVLAF 1120
Qy 1230 AKAQIQLIFFRNLNLITLLGLHLGLVFLPVLISYVG 1266
Db 1121 AQSKIFDVFYFRWFTLIIIVAAHLHLLFLPALLSLFG 1157
```

```
RESULT 5
US-09-462-136-9
; Sequence 9, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1296
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-462-136-9
```

Query Match 15.6%; Score 1078; DB 2; Length 1296;  
Best Local Similarity 24.0%; Pred. No. 2.8e-92;  
Matches 321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;

QY 76 KICPRLVTGNTQACCSAKQVLSLEASLSITKALLTRCPACSNFVNLHCHNTCSNQSL 135  
DB 63 EFCPHLLTGDN-KLCTPQAEGLTKQIAQARHLGRCPSCFDFNFAKLWCEFTCSNQD 121  
QY 136 FINVTRVAQL--GAGOLPAVVAEAF-----YQHS--FAEQSYDSCSRVRPAAATLAVG 186  
DB 122 FVISEMKPIEKEGTEPQPAEAYVNTVEYLSLTDFAEGMFSSCKDVTFCGQALRV- 180  
QY 187 TMCVGVYSALCNAQRMWFGDGTGNGL-APLDTIFHLLPEQAVGSIQIP-LNEGVARCN 244  
DB 181 -MC---TSTPCTLTNMLEFTQNLDTNPIHTKFLLYDPKTPPSDRSTYMNWNTGCD 236  
QY 245 ESQGDVATSCODCAACPAIARPOALDSTFVLGQW-----PGSLVLIILLCVPAVVT 299  
DB 237 KSARVGNPACSTSEC--NKEYANLIDLDDGKTSQGTNCNVHGIACLNIFVMLAFIGSLAV 294  
QY 300 ILLVGF-----RVAPARDKSKWDPKKGTSLSKLSKLSFSTHTLLGQFF 341  
DB 295 LLCVGFVTSYDEYTNLRQTQSGEESPXNRK-RTGAWI-----HNFVWENNA 342  
QY 342 QGHTVWASWPLTILVLSVIPVVALAAGLVFTTELTPDVELMSAPNSQARSEKAPHDOHF 401  
DB 343 RDIGMAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVVDMWSSPSRARQEBWVFNAF 402  
QY 402 GPFRTNVLTPAPRSSRYSDSLILGPKNPSGILDLDDLELLELEQLERLHLQWSPBA 461  
DB 403 GRPQRYQQIMLL--SHRDFQSGKLYGP-----VFHKOIFEELEFDILNAIKNISTQSDG 455  
QY 462 QRNISQDICYAPLNPDNTSLYDCCINSLLQYFONNRTLLLTANOT-----L 509  
DB 456 -RTITLDDVCYRPMGPG-----YDCLMSTNTYFQGNKEHLDKMSKEETVEDDDADFYF 510  
QY 510 MGQTSQVDWKNHPLYCANAPLTFFKOGTALALSCADYGPVFPFLAIGGYKGDYSEABA 569  
DB 511 SSEATTDEWNHMAACIDQPMQSK--TKSGLSCWGTGGPSAPNM-VFGKSNSTNHQAANS 567  
QY 570 LMTFSLNYPAGDPPLAQAKLWEBAFLEEMAFQRRMAGMFOVTTAERSLEDEINRTT 629  
DB 568 IMMTILVTO--RTEPIQIAELWEKEFLFKCEYREKSPKI--FSFMAERSITDEIENDA 624  
QY 630 AEDLPITFATSYIVIFYLSLALGSY-----SSMSRVMDSKATLGLGVAVLGAVMAAM 684  
DB 625 KDEIVTVVIALAFLGYVTFSLORYFVCENQLWS-ILVHSR----- 664  
QY 685 GFYSYLGRSSLLVQWVPLVLSVGDANIFIFVLEYQRLPRRPGEPREVHIGRALGRVA 744  
DB 665 GF----- 666  
QY 745 PSMLLCSLSEATCFPLGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKQBA 804  
DB 667 -----TDLPAIRTFCLAGLAVLIDVWLHCTIFLALFVMDTQRELN 707  
QY 805 SRLDVCCVKPQLPPPGQEGI-----LLG-----FQKAYAPFLHWNTRGVV 849  
DB 708 G-----KP-EFFPPYQIKLLGALYLGROATDTFTQFFHFQVAPFLHMRTRIT 758  
QY 850 LLLFLALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLNLRYFEVGPVYFVTTLGYN 909  
DB 759 GIIFTASFTITLSSKISVGFQDSMAFTEKYSISTHFRYLQKFDVGDVGPVFTVDGELD 818  
QY 910 FSEAGMNAICSSAGCNNSFTQKIQTAYATEFPEQSYLAIPASSWVDDFDIMWT-PSSCCR 968  
DB 819 WRPDPYONKCFTPPGSDTSFGNIMNYAVGHTQYTLSGEMYNWIDNYLEWISRKSPCK 878  
QY 969 LYISGNKDKCFESTVNSL-----NCLKNCMSITMGSVRSVSEQFKYLPFLNDRPNK 1024  
DB 879 VYVHPDN--TFCSTNRNKSALDDKACRTCMDF----- 908

1025 PKGGLAAYSTSVNLTSQGVLASRRFMAHYKPL--KNSQDYTEALRAARELAANITADLRK 1082  
DB 909 --DGRASPKDAISFTSGRIQASQFMTFHKCLSGSNSDFIKAMDRTARVWVSRRLERSI-- 964  
QY 1083 VFGTDPAFVFPYTIITNVFEYQYLITLPEGLFMLSCLVPTFAVSCLLIGLDLRSGLNL 1142  
DB 965 ----DDTAHVFAYSKIFPFYEQYSTIMPILTTQLFTIVVGVFGIICVTLGIDYKGAACAV 1020  
QY 1143 LSTVMILVDTGVFMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAK 1202  
DB 1021 ICOVS-----NYFHVSS--GILIEFSVNVKGYACSLRQAKDRAE 1060  
QY 1203 EATISNGSAVFAVAMTNLPGILVLGLAKAQLIQIFFFRLNLITLLGLLHGLVFLPVIL 1262  
DB 1061 STVSGISPIILSGPVVVMAGSTMFLSGAHLQIITVYFFKFLITIVSSAVHALIILPILL 1120  
QY 1263 SYVGPDVNPALALEQKBAEAAVAVMVASCPNHPSPRSVSTADNIYV-----HSFEGS 1314  
DB 1121 AFGSGRSGHSGSETSTNDNDEQHDACVLS--PTAESHSINVEEGILNRPSLLDASHILDPL 1178  
QY 1315 IKGAGAINFLPNNGROF 1332  
DB 1179 LKAEGLDKAI---GRDF 1193

RESULT 6  
US-08-540-406-19  
; Sequence 19, Application US/08540406  
; Patent No. 5837538  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; APPLICANT: GOODRICH, LISA V  
; APPLICANT: JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/540,406  
; FILING DATE: 06-OCT-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: a60190-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-540-406-19

Query Match 9.4%; Score 651.5; DB 1; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATVCSQDCAACPAIARPOALDSTFYLQMPGS 284  
DB 225 PQAVGSGIQPLNEGVARCNESQGDVATVCSQDCAACPAIARPOALDSTFYLQMPGS 284

```
Db 24 PGRPAGGRRRTGGLRRAAAPDRD-----YLRPSYCDAAFALEQI--- 65
Qy 285 LVLIILCSVFAVTVILVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAPORLLFKLGCYIQNC 98
Qy 338 GQFFQGWGTWASWPLTILVLSVIPVLAAGLVFTELTDPVELWASPNQARSEKAFH 397
Db 99 GRF-----LVUGLLIFGAFVGLKKAANLETNVEELVVEVGGVRVSRNLT 143
Qy 398 DQHFQFFRTN-QVILTAPNRSSYRDSLLGPKNFSGILDLDLLELLE--LQERLRLH 454
Db 144 RQKICEAMFNPLMIQTPKEBG-----ANVLITTEALLQHLDSALQASRVHV 190
Qy 455 QWSPQARQNSISLQDICVAP-----LNPDTNLSYDCINSLLQVFNQNRLLTLTA 505
Db 191 YMYN-----RQWKLEHLCYKSGELITETGYMDQIIEYLPCLIIITPLDCFWEGAKLQSGTA 246
Qy 506 NOTLMO-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFNFDPLEFLEELKKINYQVDSWEEMLNKAEGVGHYMDRCLNPADPDC 304
Qy 528 -----APL-----TPKOGT-----ALALSCMA 544
Db 305 PATAPNKNSKPLDMALVNGGCHGLSRKYMHWQBELIVGGTVKNSTGKLVSAHALQTMF 364
Qy 545 DYGAIVFPFLATCGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAOKLWEEAFLEMRQAF 604
Db 365 QLMTPKQMY--EHPFKGEY-----VSHINWNE---DKAAAILLEAWORTYVEVHVOSV 411
Qy 605 RMACGMFOVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSYSSWSRMVDS 664
Db 412 AQNSTQKVLSTFT--TLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
Qy 665 KATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGDDVFLAHAFSET 525
Qy 725 PRPQEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLGLAVILDF 784
Db 526 GQNKRIPEDRFTGECIKTKTGASVALTISNVTAFPMALIPALRAFSLQAAVVVNF 585
Qy 785 LIQMSAFVALLSLDSKROEASRLDVCCC-----VKPOEL-----PPG 822
Db 586 ANVLLIFPAILSMIDLYREDRLDIFCCFTSPCVSRVLOVFPQATDTHDTRYSPPPY 645
Qy 823 QEGU----- 827
Db 646 SSHSPAHEQTQITMQSTVQLRTEYDPHTHYTTAPRSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFQKAVAPFLHWTGTVVLLPLALFGVSL 861
Db 706 SSTRDLLSQFSSSLHCLBPPCTKWTLSFPAKHYPAPFLPKAKVAVVIFLGLLGVSL 765
Qy 862 YSMCHISVGLDQELALPKDSLVDLFLFLNRYFVFGAPVYFVTTLGYNFSEAGNNAICS 921
Db 766 YGTTVRDGLDITDIVPRETREYDFIAAQKYFSP-----YMW----- 803
Qy 922 SAGCNFSFTQIKIYAT-----BPPQSYLAIPASS-----WDDFDIMJ----- 961
Db 804 -----YIVTKADYPNIQHLLYDLIHRFSFNKYVMLEENKQLPKMWLHYFRDMLQGLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKFCPSTVNSLNLCKNCSITWGS 1000
Db 858 AFDSWETGKMPNPNKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFHXYL--PMFLND-----RPNKCPKGGLAAY--STSVMNLTSQDQ 1043
Db 914 INPSA--FYIYLTAWVSNDPVAYAAASQANIRPHRPWVHDK---ADYMPETRLRIPAAEP 968
Qy 1044 VLASFEMAYHKLKNSQDYTAALRAARELANITA-DLRKVGTPDPAPEVFPYITNVFY 1102
Db 969 IEYAQFPFLNGLRDTDFVEAIEKVKRTICSNYTSLSGLSSYPNG-----YFF-----LFW 1018
```

---

```
Qy 1103 EYLTILPEGLFMLSCLVETFAVSCULLGLDLRSLNLLSIWMILVDTVGFMAWDIS 1162
Db 1019 EYTLGRHLLFISVVLACTFLVCAVFLNPMPTAGII-VNVALMTVELFGMMGLIGIK 1077
Qy 1163 YNAVSLNLVSAGMSVEFVSHITRSE--AISTKPTWLERAKEATISMGSAVPAVAMTN 1220
Db 1078 LSAVPVILASVIGVEFTVHVVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
Qy 1221 LPGILVGLAKAQIQTIFFRLLNLLITLGLHLGLVFLPVLISYVG--PDVNPALALEQ- 1277
Db 1134 LGLVLMAGSEFDFIVRYFAVLAAILTILGVLNGLVLLPVLSPFGVPEVSPANGNLRL 1193
Qy 1278 --KRAEAAVAVMVASCNHPSPRVSTADNIYVNHVSFECSIKG 1317
Db 1194 PTPSPEPPPVVRPAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 7
US-08-656-055-19
; Sequence 19, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PGAVSGGIQPLNEGVARCNESQDDVATCQDCAACSCPAIARPOALDSTFVLGQMPGS 284
Db 24 PGRPAGGRRRTGGLRRAAAPDRD-----YLRPSYCDAAFALEQI--- 65
Qy 285 LVLIILCSVFAVTVILVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
```

```
Db 66 -----SKGATGRKAPLWLRKAFORLLFLKGCYIQKNC 98
Qy 338 GQFFQGGTGWASWPLTILVSVIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFH 397
Db 99 GKF-----LVVGLLIFGAFVGLKAANLETNVEELVVEGGRVSRRELYT 143
Qy 398 QHFGGFFRTN-QVILTAPNRRSYRYSLLGPNFSGIIDLLELLELLE--LOERLRLH 454
Db 144 RQKIGSEAMFNQIMQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
Qy 455 QWSPAEQRNISLQDCIYAP-----LNPDTSLYDCCINSLLQYQNNETALLLTA 505
Db 191 YMN-----RQWKLHLCYKSGELITETGMDQIIEYLPCLITPLDCFWEGAKLQSGTA 246
Qy 506 NOTLMGO-----TSQVMDKDFLY-----CAN----- 527
Db 247 --YLLGKPLRWTFPDLPLEELKKINQVDSWEEMLNKAEGVGHGMDRCPCLNPADPDC 304
Qy 528 -----APL-----TPKQGT-----ALALSQWA 544
Db 305 PATAPNKNSTKPLDMALVNLGGCHGLSRKTMHQEBELIVGTVKSTGKLVSAHALQTMF 364
Qy 545 DYCAPVFPFLAIGYKGDYSEAEALIMTFLSNYPAGDPRLAQAQKLWEEAFLEEMRAFQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINMNE---DKAAAILBEAMQRTYVVEVHQSV 411
Qy 605 RRMAGNFQVTFTAERSLEDEINRTTAEDLPIFATSVIVIFLYISLALGSYSSMSRVWVDS 664
Db 412 AQNSTQKVSFTT-TWLDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS 465
Qy 665 KATLGGGVAVLGAAMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIFIVLEBYQRL 724
Db 466 QCAVGLAGVLLVALSVAAGLGLSGLIGISFNAAATQVLPFLALGVGDVDFLLAHAFSE 525
Qy 725 PRPGCPREVIHGRALGRVAPSMCLSLSEALCFPLGALTPPAVTFALTSLGLVILDF 784
Db 526 GQNKRIPFEDRTGECUKRTGASVALTISINVTAFMAALIPALRAFSLQAAVVVVF 585
Qy 785 LQMSAFVALLSLSKQBSASRLDVCCC-----VKQOEL-----PPPG 822
Db 586 AMVLLIFPALLSMDLYREDRLDIFCCFTSPCSRVIOVEQAYTHTDNTREYSPFPY 545
Qy 823 QSEGL----- 827
Db 646 SSHSFAHQITMQSTVQLRTEYDPTHVYVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFFQKAYAPFLHWTITRCVVLVLLFLALFGVSL 861
Db 706 SSTRDLLSFDSLSHCLBPCTKWTLSFAEKHYAPFLPKAKVVIPLFLGLLGNVL 765
Qy 862 YSMCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGPVVFVTLGYNFSSEAGMNAICS 921
Db 766 YGTRVRDGLDITDIVPRETREVDFIAAQKFSP-----YNM----- 803
Qy 922 SAGCNNFSTQKIQYAT-----EFPEQSYLEIPASS-----WVDDFDIDLW----- 961
Db 804 -----YIVTQADYPNIQHLLYDLHRSFNSVKYVMLEENKQLPKMWLHYFRDLWQLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKCPTVNSLNCNCKMSITMGS 1000
Db 858 AFDSOWETSKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFHKL-PWFLND-----RPNIKCPKGLAAY--STSNNLTSDGQ 1043
Db 914 INPSA--FYIYLTAWNSDPVAYASQANIRPHRPEWHDK---ADYMPETRLRIPAAP 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKKVFGTDPAPFVFFYTTINVPY 1102
Db 969 IEVAQPPFYNLGRDTSDFVEAIEKVRTICSNVTSIGLSSYPNG-----YFP---LFW 1018
Qy 1103 EQYLTPLPGLFMLSCLVPTTAVSCLLGLDLRSLGNLLSIVMLVTVTGFMALWDIS 1162
Db 1019 EQYIGLRHWWLLFISVVLACTFLVCAVFLNPNWTAGII-VMVLAJMTVFLFGMMGLIGIK 1077
```

---

```
Qy 1163 YNAVSLINLSAVGMSVEFVSHITRSF--AISTKPTWLRKAEATTISMGSAVPAVAMTN 1220
Db 1078 LSAVPVILLIASVIGVETTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
Qy 1221 LGGILVLGAQAQKLIQIFFRNLNLLITLGLHLGLVFLPVILSYVG--PDVNPALALEQ- 1277
Db 1134 LLGLVLMAGSEDFIVRYFFFAVLAILTILGLVNLGLVLLPVLLSFFGPFYEVSPANGINRL 1193
Qy 1278 --KRAEEAAVAVVWASCPNHPRSVSTADNIYVNHSPGSIKG 1317
Db 1194 PTPSPBPSPVVRFAFPMPGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 8
US-08-954-668-19
; Sequence 19, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Foley, Hoag & Elliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-668-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PQQAVSGGIQPLNEGVARCNESQGDVATCSQDCAASCAPARPAALDSTFYLQGMPS 284
Db 24 PCRAGGRRRTTGLRAAAPDRD-----YLHRRPSYCDAAAFALQI--- 65
Qy 285 LVLIILISVFAVWITLLVGFVAVAPARSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGATGRKAPLWLRKAFORLLFLKGCYIQKNC 98
Qy 338 GQFFQGGTGWASWPLTILVSVIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFH 397
Db 99 GKF-----LVVGLLIFGAFVGLKAANLETNVEELVVEGGRVSRRELYT 143
Qy 398 QHFGGFFRTN-QVILTAPNRRSYRYSLLGPNFSGIIDLLELLELLE--LOERLRLH 454
Db 144 RQKIGSEAMFNQIMQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
Qy 455 QWSPAEQRNISLQDCIYAP-----LNPDTSLYDCCINSLLQYQNNETALLLTA 505
Db 191 YMN-----RQWKLHLCYKSGELITETGMDQIIEYLPCLITPLDCFWEGAKLQSGTA 246
Qy 506 NOTLMGO-----TSQVMDKDFLY-----CAN----- 527
Db 247 --YLLGKPLRWTFPDLPLEELKKINQVDSWEEMLNKAEGVGHGMDRCPCLNPADPDC 304
Qy 528 -----APL-----TPKQGT-----ALALSQWA 544
Db 305 PATAPNKNSTKPLDMALVNLGGCHGLSRKTMHQEBELIVGTVKSTGKLVSAHALQTMF 364
Qy 545 DYCAPVFPFLAIGYKGDYSEAEALIMTFLSNYPAGDPRLAQAQKLWEEAFLEEMRAFQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINMNE---DKAAAILBEAMQRTYVVEVHQSV 411
Qy 605 RRMAGNFQVTFTAERSLEDEINRTTAEDLPIFATSVIVIFLYISLALGSYSSMSRVWVDS 664
Db 412 AQNSTQKVSFTT-TWLDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS 465
Qy 665 KATLGGGVAVLGAAMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIFIVLEBYQRL 724
Db 466 QCAVGLAGVLLVALSVAAGLGLSGLIGISFNAAATQVLPFLALGVGDVDFLLAHAFSE 525
Qy 725 PRPGCPREVIHGRALGRVAPSMCLSLSEALCFPLGALTPPAVTFALTSLGLVILDF 784
Db 526 GQNKRIPFEDRTGECUKRTGASVALTISINVTAFMAALIPALRAFSLQAAVVVVF 585
Qy 785 LQMSAFVALLSLSKQBSASRLDVCCC-----VKQOEL-----PPPG 822
Db 586 AMVLLIFPALLSMDLYREDRLDIFCCFTSPCSRVIOVEQAYTHTDNTREYSPFPY 545
Qy 823 QSEGL----- 827
Db 646 SSHSFAHQITMQSTVQLRTEYDPTHVYVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFFQKAYAPFLHWTITRCVVLVLLFLALFGVSL 861
Db 706 SSTRDLLSFDSLSHCLBPCTKWTLSFAEKHYAPFLPKAKVVIPLFLGLLGNVL 765
Qy 862 YSMCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGPVVFVTLGYNFSSEAGMNAICS 921
Db 766 YGTRVRDGLDITDIVPRETREVDFIAAQKFSP-----YNM----- 803
Qy 922 SAGCNNFSTQKIQYAT-----EFPEQSYLEIPASS-----WVDDFDIDLW----- 961
Db 804 -----YIVTQADYPNIQHLLYDLHRSFNSVKYVMLEENKQLPKMWLHYFRDLWQLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKCPTVNSLNCNCKMSITMGS 1000
Db 858 AFDSOWETSKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFHKL-PWFLND-----RPNIKCPKGLAAY--STSNNLTSDGQ 1043
Db 914 INPSA--FYIYLTAWNSDPVAYASQANIRPHRPEWHDK---ADYMPETRLRIPAAP 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKKVFGTDPAPFVFFYTTINVPY 1102
Db 969 IEVAQPPFYNLGRDTSDFVEAIEKVRTICSNVTSIGLSSYPNG-----YFP---LFW 1018
Qy 1103 EQYLTPLPGLFMLSCLVPTTAVSCLLGLDLRSLGNLLSIVMLVTVTGFMALWDIS 1162
Db 1019 EQYIGLRHWWLLFISVVLACTFLVCAVFLNPNWTAGII-VMVLAJMTVFLFGMMGLIGIK 1077
```

Db 144 RQKIGBEAMFNQMIQTPEKG-----ANVTTEALLQHLDSALQASRVHV 190  
QY 455 QVMSPEAQRNLSLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505  
Db 191 YMYN----RQWLEHLCKYSGELITETGYMDQIIIEVLYPCLIIITPLDCFWEGAKLSGTA 246  
QY 506 NQTLMO-----TSQVDMKDHFLY-----CAN----- 527  
Db 247 --YLLGKPLRWTFNDFPLEFLEELKKNYQVDSWEEMLNKARVGHGYMDRCLNADPDC 304  
QY 528 -----APL-----TFKQGT-----ALALSMA 544  
Db 305 PATAPKNSTKPLDMALVNGGCHGLSRKYMHWQBELVGTGVKNSTGKLSAHALQTMF 364  
QY 545 DYGAIVFPFLAIGGYKGYSEAEALIMTFLSNYPAGDPRLAQAALWEEAFLEEMRAFO 604  
Db 365 QLMTPKQMY---EHRFGYEV-----VSHINWNE-----DKAAALILEAWQRTYVEVHVQSV 411  
QY 605 RRMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSSWSRVMVDS 664  
Db 412 AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465  
QY 665 KATLGGVAVVYLGAVMAAMGFSSYLGIRSSVLVQVPPFLVLSYGADNIFIFVLEYQRL 724  
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLAHAFSET 525  
QY 725 PRPGEPREVIHGRALGRVAPSMMLCSLSEACFFLPGALTPMPAVRTFALTSLGLAVILDF 784  
Db 526 GQNKRIPEPDRGTGCKRTGASVALTSISNVTAFFMAALIPALRAFSLQAAVVVVFNF 585  
QY 785 LIQMSAFVALLSLDKROEASRLDVCCC-----VKPOEL-----PPG 822  
Db 586 ANVLIFPALLSMDLYRREDRLDIFCCFTSPCSRVIQVEPQATYTHDNTRISPPPPY 645  
QY 823 QGEG----- 827  
Db 646 SSHSPAHEQTITMQSTVOLRTEYDPHTHYVYTTAEPRESEISQPVVTQDTLSQSPST 705  
QY 828 -----LLGPFQKAYAPLHWTIRGVVLLFLALFGVSL 861  
Db 706 SSTRDLLQFSDSSLHLCLEPPCKWTLSSEFAKHYAPFLPKAKVQVFIPLGLLGVSL 765  
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAQVYFVTTLGYNFSSEAGNNAICS 921  
Db 766 YGTRVRDGLDITDIVPRETREYDFIAAQFYFSP-----YMW----- 803  
QY 922 SAGCNFSTQKIQYAT-----EFPEQSYLAIPASS-----WVDDFIDWL----- 961  
Db 804 -----YIVTOKADYPNIQHLLYDLHRSFSNVKYVMLEENKQLPKQWMLHYFRDMLQGLQD 857  
QY 962 -----TPSS-----CCRLXISGPNKPCPSTVNSLNCNLCNCSITMGS 1000  
Db 858 AFDSWETGKMPNNYKNGSDGVLAYKLLVOTGSRDK--PIDISQLT--KORLVADAGI 913  
QY 1001 VRPSVEQFHKYL--PWFAND-----RENKPCPKGLAAY--STSVALTSDGQ 1043  
Db 914 INPSA--FVIYLTAVNSNDPVAYAASQANIRPHREWVHDK--ADYMPETRLAIPAEP 968  
QY 1044 VLASFPMAYHKLKNSQDYTEALRAELANITA-DLRKVPGTDPAPFEPYITNVFY 1102  
Db 969 IEYAQFPFLNGLRDTSDFVEAIEKVRTICSNYTSLSGLSSYPNG-----YFF-----LFW 1018  
QY 1103 EOYLTPLEGFLMLSLCVPTFAVSCLLGLDLSRGLNLLSIVMVLVDTVGFMAWDIS 1162  
Db 1019 EOYIGLRHLLLFISVVLACTFLVCVAFLLNPWTAGII--VWVLMVTVELFGMGLGIK 1077  
QY 1163 YNAVSLINLVAGMSVEFVSHITRSP--AISTKPTWLERAKEATISMSGSAVFAGVAMTN 1220  
Db 1078 LSAVPWVILIASVGIGVEFTVHVALAFLTAIGDKN---RRAVLALEHMFAPVLDG-AVST 1133  
QY 1221 LPIGLVILGAKAQLIQIIEFFRNLNLLITLLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277

Db 1134 LLGVMLAGSEDFIVRYFFAVLAILTILGVNLGLVLLPVLLSFFPGYPPEVSPANGLNRL 1193  
QY 1278 --KRAEAVAAVMVASCNHPSPRVSTADNIYVNHVSFEGSIKG 1317  
Db 1194 PTPSPPEPPSVRFPAMPFGHTH--SGSDSDSDSEYSSQTTVSG 1233  
RESULT 9  
US-09-268-140-5  
; Sequence 5, Application US/09268140  
; Patent No. 6268176  
; GENERAL INFORMATION:  
; APPLICANT: Gemmill, Robert M.  
; APPLICANT: Drabkin, Harry A.  
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED  
; FILE REFERENCE: 93445-00004  
; CURRENT APPLICATION NUMBER: US/09/268,140  
; CURRENT FILING DATE: 2000-03-12  
; PRIOR APPLICATION NUMBER: US 60/077,723  
; PRIOR FILING DATE: 1998-03-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1447  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-268-140-5  
Query Match 9.4%; Score 651.5; DB 2; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 2460; Indels 421; Gaps 47;  
QY 225 PGQAVSGGIQPLNEGVARCNESQGDVATCSCQDCAACPAIARPOALDSTFFYLQMPGS 284  
Db 24 PGRPAGGRRRTTGLRRAAAPDRD-----YLRPSYCDAAFALEQI--- 65  
QY 285 LVLLIILCSVFAVVTILLVGRVAPARADSKMVDPK-----KGTSLSDKLSFSTHTLL 337  
Db 66 -----SKGATGRKAPLWLRKAFQFELLFLKCYIQKNC 98  
QY 338 GQFQFGGTWVANSWPLTILVLSVPVVALAAGLVFTETDTPVELMSAPNSQARSEKAPH 397  
Db 99 GKF-----LVVGLLIIFGAFVAGLKAANLETNVEELWVEVGGVRSRELYNT 143  
QY 398 DQHFPGPPFTN-QVILTPAPNRSSRYVDSLGLPKNFGSGLDLDLLELLE--LQERLRL 454  
Db 144 RQKIGBEAMFNQMIQTPEKG-----ANVTTEALLQHLDSALQASRVHV 190  
QY 455 QVMSPEAQRNLSLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505  
Db 191 YMYN----RQWLEHLCKYSGELITETGYMDQIIIEVLYPCLIIITPLDCFWEGAKLSGTA 246  
QY 506 NQTLMO-----TSQVDMKDHFLY-----CAN----- 527  
Db 247 --YLLGKPLRWTFNDFPLEFLEELKKNYQVDSWEEMLNKARVGHGYMDRCLNADPDC 304  
QY 528 -----APL-----TFKQGT-----ALALSMA 544  
Db 305 PATAPKNSTKPLDMALVNGGCHGLSRKYMHWQBELVGTGVKNSTGKLSAHALQTMF 364  
QY 545 DYGAIVFPFLAIGGYKGYSEAEALIMTFLSNYPAGDPRLAQAALWEEAFLEEMRAFO 604  
Db 365 QLMTPKQMY---EHRFGYEV-----VSHINWNE-----DKAAALILEAWQRTYVEVHVQSV 411  
QY 605 RRMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSSWSRVMVDS 664  
Db 412 AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465  
QY 665 KATLGGVAVVYLGAVMAAMGFSSYLGIRSSVLVQVPPFLVLSYGADNIFIFVLEYQRL 724  
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLAHAFSET 525  
QY 725 PRPGEPREVIHGRALGRVAPSMMLCSLSEACFFLPGALTPMPAVRTFALTSLGLAVILDF 784

Db 526 QNKRIPFDETECLKRTGASVALTSISNVTAFMAALIPFALRAFSLOAAVVFNF 585  
Qy 785 LLQMSAFVALLSDSKRQASRLDVCC-----VKPQEL-----PPG 822  
Db 586 AMVLLIFPALLSMDLYRERDRDLDFCCFSPCVSRVQVPEQAYTTHDNTRYSPPPY 645  
Qy 823 QSGGL----- 827  
Db 646 SSHSFAHETQITMQSTVQLRTEYDPHTHYVYTTAEPRSEISVQPVVTQTLLSCQSPST 705  
Qy 828 -----LLGFFQKAYAPFLHWTIRGVLLHFLALFGVSL 861  
Db 706 SSTRDLLSQSDSLHCLPECTKWTLSFAEKHYAPFLKPKAKVWVIFLFLGLLGVSL 765  
Qy 862 YSMCHISVGLQBELALPKDSYLDYFLFLNRYEVEGAPVYVTTLLYNFSSSEAGMAICS 921  
Db 766 YGTRVRDGLDLDIVPRETREDYFIAAQKYFSF-----YNM----- 803  
Qy 922 SAGCNFPSTQKIQYAT-----EPPEQSYLAIPASS-----WVDDFIDWL----- 961  
Db 804 -----YIVTKADYPNIQHLLYDLHRSFNSVYVMLEENKQPKWMLHYFRDMLQGLQD 857  
Qy 962 -----TPSS-----CCRLYISGNKDKPCPSTVNSINCLKNCSITMGS 1000  
Db 858 AFDSDWETGKIMPNNYKNGSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913  
Qy 1001 VRPSVEQFHKYL--PWFLND-----RPNIKCPKGLAAY--STSVNLTSDGQ 1043  
Db 914 INPSA--FYIYLTAWNSNDVAVASQANIRPHRPEWHDK---ADYMPETRURIPAAEP 968  
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA--DLRKVPQGTDPAFEPYPTIYNVY 1102  
Db 969 IEVAQFPFYNGLRDTSDFVEATEKVRTICSNVTSGLSSYPNG-----YPP-----LFW 1018  
Qy 1103 EQVLTTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMAWDIS 1162  
Db 1019 EYQIGURHLLGLFISVLACTFLVCAVFLNPNWTAGII--VMVLALMTVELFGMGLIGIK 1077  
Qy 1163 YNAVSLINLVSAVGMSEVFSHTSRF--AISTKPTWLBRAKEATISMGSAVPAGVAMTN 1220  
Db 1078 LSAVPVILIASGVIGVETVHVHVAFLTAIGDNK---RRVLALEHMFAPVLDG-AVST 1133  
Qy 1221 LFGILVLGAKAQLIQIPFRLMLLTLGLLHGLVFLVILSYVG--PDVNPALALEQ- 1277  
Db 1134 LLGVLMAGSEDFIVRYFFAVLAULTILGVNLGLVLLPVLSSFFGYPYEVSPANGINRL 1193  
Qy 1278 --KRAEAVAAVMVASCNHPSPSVSTADNIYNHSEGSIGK 1317  
Db 1194 PTPSPPEPFSVVRFAFMPGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 10

US-08-918-658-19  
; Sequence 19, Application US/08918658  
; Patent No. 6429354  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; GODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/918,658  
FILING DATE: 22-Aug-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/656,055  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/540,406  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: a60190-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1447 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-08-918-658-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;  
Qy 225 PQQAVSGSIQPLNEGVARCNESQGGDVATCSQDCAASCPAIPARQALDSTFYLQMPGS 284  
Db 24 PGRAGGRRRTTGLRRAAAPDRD-----YLHRPSYCDAAFALEQI--- 65  
Qy 285 LVLIILCSVFAVITLLVFRVAPARDKSKWDPK-----KGTSLSDKLSFSHTLL 337  
Db 66 -----SKGKATGRKAPLWRAKFORLLFKLGCYIQKNC 98  
Qy 338 GQFFQGWGTWASWPLTILVSVIPVVAALAGVFTLTDPVELWSAPNSQARSEKAPH 397  
Db 99 GKF-----LVVGLLIFGAFVAGLKAANLETNVEELVGVGVRSELNT 143  
Qy 398 DQHFQFPFRTN-QVILTAPNRSSYRYSLLLGPKNPSGILDLDLLELE--LQERLRLH 454  
Db 144 RQKIGBEAMFNQMIQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190  
Qy 455 QVWSPQARQNISLQICVAP-----LNPDTSLYDCINSLLQYFQNNRTLLLLTA 505  
Db 191 YMYN---RQWKLEHLCYKSGELITETGYMDQIIIEYLPCLITPLDPCFWEQAKLQSGTA 246  
Qy 506 NQTLMGQ-----TSQVDMKDHPLY-----CAN----- 527  
Db 247 --YLLGKPLRWTNFDPLEFLEELKKINYQVDSWEELNKAIEVGHYMDRCPCLNPADPC 304  
Qy 528 -----APL-----TFKDG-----ALALSCMA 544  
Db 305 PATAPKNSTKPDMLVNLGCGCHGLSKRYMHWQBELIVGTVKSTGKLVSAHALQTFP 364  
Qy 545 DYGAVPFPFLATGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAELEENRPFQ 604  
Db 365 QLMTPEKQMY---EHFKGYEY-----VSHINWNE---DKAAAILAEAWQRTVYVHVQSV 411  
Qy 605 REMAGMFWQVTFTAERSLDEINRTTAEDLPFATSVYIVFLYISLALGYSYSSWSRVWDS 664  
Db 412 AQNSTQKVLSPFT--TLLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465  
Qy 665 KATLGLGVAVVLGAVMAAMGFYSYLGIRSSILVILQVVPFLVLSYADNIFIFLVEYQRL 724  
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGVDDVFLLAHAFST 525  
Qy 725 PRPQPEPREVHIGRALGRVAPSMMLCSLSEACFPLGALTTPMPAVRTTALTSGLAVIDLF 784

526 GONKRIPEPTDRTGECLKRTKRTGASVALTISINVTAFPMALIPALRAFSLQAAVVVNF 585  
785 LLOMSAFVALLSLDKROEASRLDVCCC-----VKPOEL-----PPG 822  
586 AMVLLIFPAILSMDLYREDRLDIFCCFTSPCVSRVQVFEQAYTTHDTRYSPPPY 645  
823 QGEGL----- 827  
646 SSHSFAHETQITMSTVOLRTYEDPHYVYTTAEPSEISVQPVTVTQDTLSQSPST 705  
828 -----LLGPFQKAYAPFLHWTIRGVVLLFLALFGVSL 861  
706 SSTRLLSQFSSSLHCLPEPPCTKWTLSFAEKHYAPFLPKAKVVVIFLGLGVSL 765  
862 YSMCHISVGLDELALPKDSVLLDYFLFNRYFEVGAPVYFVTILGYNFSSEAGNVALCS 921  
766 YGTRVRDGLDITDVPRETREYDFIAAQPKYFSF-----YNN----- 803  
922 SAGCNFSTOKIOYAT-----EFPEQSILAI PASS-----WVDDFIDWL----- 961  
804 -----YIVTQADYPNIQHLLYDLHRSFNVKVMLEENKQLPKWMLHYFRDMLQGLD 857  
962 -----TPSS-----CRLYISGNKKKFCPSTVNSLNCNKMSITWGS 1000  
858 AFDSWETGKIMPNNYKNGSDGVLAYKLVQTGSRDK--PIDISQLT--KQRLVDADGI 913  
1001 VRPSVEQHKYL--PWFLND-----RPNKCPKGLAAY--STSVALTSDQG 1043  
914 INPSA--FYIYLTAWVSNPDVAYASAQANIRHPREPWHDK---ADYMPETRLRIPAAEP 968  
1044 VLASRFMAHYKPLKNSQDYTEALRAARELANITA-DLRKVFGTDPAPFVPPYITVNFY 1102  
969 IYEAQFPYINGLRDTSDFEAEIKVRTICSNYISLGLSSYPNG-----YFP-----LFW 1018  
1103 EOYLTLPEGFMLSCLVPTFVAVSCLLLGLDLSGLNLNLISVMILVDTVGFMALWDIS 1162  
1019 EOYIGLRHLLIFISVWLACTFLVCVFLNPNWTAGII-VMVLMALMTVELFGMGLIGIK 1077  
1163 YNAVSLINLVAGNSVEFVSHITSEF--AISTKPTWLERAKEATISNGSAVFAGVAMTN 1220  
1078 LSAPVSVLLIASVGIGVEFTVHVVALAFLTAIGDKN---RRAVLALERNFAPVLOG-AVST 1133  
1221 LPIIIVLGLAKAQITQIEFFRLNLITLLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277  
1134 LLGLVLMAGSEFDFIVRYFFAVLAILTILGLVINGLVLPLVLSFFGPPYEVSPANGLNRL 1193  
1278 --KRAEEAAMVAVASCPNHSRSTADNIVNHSFEKSIG 1317  
1194 PTPSPPPPVVRFAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 11

US-09-724-631-19  
; Sequence 19, Application US/09724631  
; Patent No. 6551782

GENERAL INFORMATION:

APPLICANT: SCOTT, MATHEW P  
; GOODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/724-631  
FILING DATE: 28-No. 6551782-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/656,055  
FILING DATE: 1996-05-31  
APPLICATION NUMBER: 08/540,406  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: a60190-1  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1447 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-724-631-19  
  
Query Match 9.4%; Score 651.5; DB 2; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;  
  
Qy 225 PGQAVSGGIQPLNEGVARCNESQGDVATCSCQDCAASCFAIARPAQALDSTFYLGQMPGS 284  
Db 24 PGRPAGGRRRTTGLRRAAPDRD-----YLRHPSYCDAAFALEQI--- 65  
Qy 285 LVLIILCSVFVAVVTILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHLL 337  
Db 66 -----SKGATGRKAPLMLRAKFORLLFLKGCYIQKC 98  
Qy 338 GQPFQGWGTWVSWPLTILVLSVIPVVALAAGLVFTLTDPVELASAPNSQARSEKAFH 397  
Db 99 GKP-----LVVGLLIIFGAFVAGUKAANLETNVBELVVEVGRVSRRELYNT 143  
Qy 398 DQHFPPPTFTN-QVILTAPNRSSRYVDSLILGPKNFSGIILDLILLLELLE--LQERLHL 454  
Db 144 RQKIGBEAMNPQLMIQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190  
Qy 455 QVWSPEAQRNISIQDICYAP-----LNPNTSYDCCINSLLQYFQNNRTLLLLTA 505  
Db 191 YMYN-----RQWKLEHLCKYSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA 246  
Qy 506 NOTLMGO-----TSQVDMKDHFLY-----CAN----- 527  
Db 247 --YLLGKPPPLRWTFNDFLEBELKKINYQVDSWEEMLNKAEEVGHGYMDRPLNADPDC 304  
Qy 528 -----APL-----TFKQGT-----ALALSCMA 544  
Db 305 PATAPNKNSTKPLDMALVINGCHGLSRKYMHWQEEELIVGCTVKNSTGKLSAHALQTMF 364  
Qy 545 DYGAVPFPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEEMRAFQ 604  
Db 365 QLMTPKQMY---BHFKGYEY-----VSHINWNE-----DKAAAILAEAWQRTYVEVHVQS 411  
Qy 605 RRMAGMFQVFTFAERSLEDSINRTABDLPIFATSYIVIFLYTSLAGSYSSRSRVNVD 664  
Db 412 AQNSTQKVLSTFT--TLLDDILKSFSDSVIRVAGSYLLMLAYACLTM---LRWD--CSKS 465  
Qy 665 KATLGLGGVAVLVGAVMAAMGFFSYLGRSSVLVQVVPFLVLSVGDADNIFIFVLEYQRL 724  
Db 466 QGAVGLAGVLLVALSVAAGLGLCLSLIGISFNAATQVLPFLALGVGVDDVFLAHAFSET 525  
Qy 725 PRPGEPREVHIGEALGRVAPSMLLCSLSAI CFFLGLATPMAVRVTRFALTSLGLAVILDF 784  
Db 526 GONKRIPEPTDRTGECLKRTKRTGASVALTISINVTAFPMALIPALRAFSLQAAVVVNF 585

```
QY 785 LLOMSAFVALLSLDSKQASRLDVCCC-----VKPQEL-----PPPG 822
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 586 AMVLLIFPAILSMDLRYREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHTNTRYSPPPY 645
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 823 QCEGL-----827
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 646 SSHSFAHETOITMQSTVQLRTEYDHPHVVYTTABPRSEISVQPVTVTQDTLSCQSPST 705
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 828 -----LLGFFQKAYAPFLHMITRGVLLHFLALFGVSL 861
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 706 SSTRDLLSQSDSSHLCPCTKTWLLSSFAEKHYAPFLPKPAKVWVIFLFLGLGVSL 765
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 862 YSMCHISVGLDQELALPKOSYLLDYFLFNRYFEGVAPVYVTVTLGNFSSSEAGMAICS 921
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 766 YGTRVRDGLDLTDIVPRETREYDFIAAQPKYSF-----YNN-----803
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 922 SACCNFSFQKIQYAT-----EPREQSYLAIPASS-----WVDDFDLW-----961
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 804 -----YIVTQKADYPIQHLLYDLHRSFNSVKYVMLENKQPKRWMLHYFRDWLQGLQD 857
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 962 -----TPSS-----CCRLYISGNKDKFCPTVNSLNCCLKNCMSITMGS 1000
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 858 APDSWETGKIMPNYKNGSDDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGI 913
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1001 VRPSVEQFKYL--PWFLND-----RPNKCPKGGLAAY--STSVNLTSDGQ 1043
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 INPSA--FYILTAWVSNPDVAVYASQANIRPHRPWVHDK--ADYMPETRLRIPAAEP 968
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1044 VLASRPWAYHKLKNSQDYTEALRAARELAANITA--DLRKVPCTDPAPEVFPYTTINVFY 1102
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 969 IEYAQPFYINGLRDTSDFEALFKVKTICNSYLSGLSSYPNG-----YFP-----LFW 1018
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1103 EQVLTLPBGLFMLSCLVPTFAVSCLLGLDLRSGLLNLSTVMILVDTVGFMALWDIS 1162
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1019 EQYIGLRHLLLFISVVLACTFLVCAVFLNPNWTAGII--VWVALMTVFLGMMGLIGIK 1077
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1163 YNAVSLINLVASGMSEVFSVSHTRSF--AISTKPTWLERAKEATISMGSAVPAGVAMTN 1220
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1078 LSAVPVILLIASVGIGVEFTVHVALAFLTAIGDKN--RRVLALEHMFAPVLDG-AVST 1133
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1221 LFGILVLGAKAQLIQIFPRNLNLTILGLHGLVFLVILSYVG--PDVNPALALEQ- 1277
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1134 LLGLVLMAGSEDFIVRYFAVLAITILGLVNLGLVLLFVLSFFGYPYEVSPANGLNRL 1193
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1278 --KRAEEAVAAVWASCPNHPRSVSTADNIYVNHSPSGSIKG 1317
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1194 PTPSPPEPPSVVRPAMPQGH--SGSDSSDSEYSQTTVSG 1233
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 12
US-08-954-701A-19
; Sequence 19, Application US/08954701A
; Patent No. 6610507
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-701A-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQQAVGSGIOPNEGVARCNESQGDVATCSCQDCACSPAIARPOALDSTFFVLGMPGS 284
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 PGRPAGGRRRTTGLRRAAAPDRD-----YLHRSYCDAAFALEQI--- 65
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 LVLIILCSVPFAVTVILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 -----SKGKATGRKAPLWLRRAKQRLFLKGCYQKNC 98
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 GQFFQGWGTWASWFLTILVSVIPVVALAAGLVFTETDPELVMSAPNSQARSEKAFH 397
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 GKF-----LVGCLLIFGAFVGLKAANLETNVEELWVEGVGRVRELNTY 143
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 DQHFQFPFRTN--QVILTAPNRSSYRDSLLGPKNFSGILDDLLLELLE--LOEKLRLH 454
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 ROKIGEEAMFNQPMIQTPKBEG-----ANVLTTEALIQHLDSALQASRVHV 190
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 QVWSPEARQNISLODICVAP-----LNPDNTSLVDCINSLLYQFNRRNTLLTLTA 505
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 YMN-----ROWKLEHCYKSGELITETGYMQQII EYLYPCLLIITPLDPCFWGAKQSGTA 246
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN-----527
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 --YLLGKPLRWTFNDFPLEELELKKINYQVDSWEMLNKAEGVGHGMDRCPCLNPADPC 304
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 528 -----APL-----TFKQGT-----ALALSCVA 544
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 PATAPNKNSTKPLDMALVNLGGCHLSRKYMHQBELIVGTVTKNSTGKLSAHALQTMF 364
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 DYGAVPFPPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAQ 604
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 QLMTPKQMY---EHPKGYEY-----VSHINWE-----DKAAAILQWRTTVEVVHQSV 411
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 605 RRMAGMFQVTFTAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSWSRVMD 664
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 AQNSTQKVLSTFT--TTLDDILKSPDSVSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 665 KATIGLGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGANIFIFVLEYORL 724
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLFALGVGDVDFLLAHAFSE 525
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 725 PRRPGEPREVHIGRALGVAPSMLLCSLSEALCFPLGALTMPAVRTSLGSLAVILDP 784
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 526 GONKRIPEDETRTGECKRTGASVALTISNTAFPMALIPALRAFSLQAQVAVVVFNF 585
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 785' LLQMSAFVALLSLDSKQASRLDVCCC-----VKPQEL-----PPPG 822
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 586 AMVLLIFPAILSMDLRYREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHTNTRYSPPPY 645
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 823 QCEGL-----827
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Db 646 SSHFAHETQITMSTQVLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSQSPST 705  
Qy 828 -----LLGFFQKAYAPFLHWHITRGVVLVLLFLALFGVSL 861  
Db 706 SSTRDLSQFSDSLHCLPCTKTWLTSSFAEKHYAPFLKPKAKVVVIFLFLGLLGVSL 765  
Qy 862 YSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTGLGNFSSEAGMNAICS 921  
Db 766 YGTRVRDGLDLDIVPRETREYDFIAAQFYFSF-----YNN----- 803  
Qy 922 SAGCNFSFTQKIQYAT-----EFPEQSYLAIPASS-----WVDDFDWL----- 961  
Db 804 -----YIVTQKADYPMIQLHLLYDLHRSFNSVNVVLEENKQLPKWMLHYFRDMLQGLQD 857  
Qy 962 -----TPSS-----CCRLYISGNKDKFCPSTVNSLNCNKCMSITWGS 1000  
Db 858 AFSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVDADGI 913  
Qy 1001 VRPSVEQPHKYL--PWFLND-----RPNIKCPKGGGLAAY--STSNNLTSQDQ 1043  
Db 914 INPSA--FYIYLTAVSNDFPVAYAASQANIRPHRPWVHDK--ADYMPETRLRIPAE 968  
Qy 1044 VLASRFWAYHKPLKNSQDYTEALRAARELANITA--DLRKVPGTDPAPFVPPYITITNVFY 1102  
Db 969 IEYAQFPFYLNGRLDTSDFEAEIKVRTICSNYTSGLSSYPNG-----YPP-----LFW 1018  
Qy 1103 EOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYMLVDTVGFMAWDIS 1162  
Db 1019 EOYIGLRHLLLFISVVVLAFTFLVCAVFLNPNWTAGII--VMVLAMTVLFGMGLIGIK 1077  
Qy 1163 YNAVSLINLVSAGVMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVFAGVAMTN 1220  
Db 1078 LSAVPVILLASVIGVEFTVHVALAFLTAIDCKN--RRAVLALEHMFAPVLDG-AVST 1133  
Qy 1221 LPHGLVLGAKAQLIQIIPFFRLNLITLLGLHLGVLPVLISYVG--PDVNPALALEQ- 1277  
Db 1134 LGLVLMAGSEDFIVRYFFAVLAILITLGLVNLGLVLLPVLSSFFGVPYEVSPANGLNRL 1193  
Qy 1278 --KRAEVAAMVAVASCNHRSRSTADNIYVNHISFEGSING 1317  
Db 1194 FTPSPPEPPSVVRFPAMPEGHTH--SGSDSDSDSYSSQTTVSG 1233

## RESULT 13

US-09-807-007-6  
; Sequence 6, Application US/09807007  
; Patent No. 6881833  
; GENERAL INFORMATION:  
; APPLICANT: ZAPHIROPOULOS, Peter et al.  
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY  
; FILE REFERENCE: 2921-0130P  
; CURRENT APPLICATION NUMBER: US/09/807,007  
; CURRENT FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1447  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-807-007-6

Query Match 9.4%; Score 651.5; DB 2; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;  
Qy 225 PQAVGSGIOPLNEGVARCNBSQGDVATCSQDCAASCPAIAARPQALDSTFYLGMPGS 284  
Db 24 FGRPAGGRRRTTGLRRAAPDRD-----YLRPSYCDAAFALEQI--- 65  
Qy 285 LVLIILCSVAVNTVILLVGFVAPARDKSNVDPK-----KGTSLSDKLSFSTWLL 337  
Db 66 -----SKGATGKRAPLWLRKAFQRLFLKGLCYIQKNC 98

Qy 338 GQFFQCGMTWVASWPLTILVLSVIPVALAAGLVFTTELTPDVELMSAPNSQARSEKAFH 397  
Db 99 GKF-----LVVGLLIIFGAFVAGLKAANLETNVELWVEVCGVRSRELNYT 143  
Qy 398 DQHFQGFPPFTN-QVILTAPNRSRYDSLLGPKNFSGLDLDLLELLE--LQERLRLH 454  
Db 144 RQKIGEEAMENPQMIQTQKEG-----ANVLTTEALLQHLDSALQASRVHV 190  
Qy 455 QVWSPQAQRNISLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505  
Db 191 YMYN-----RQWLEHLCYKSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA 246  
Qy 506 NQTLMGQ-----TSQVMDKDHFLY-----CAN----- 527  
Db 247 --YLLGKPPPLRWTFPDLPLEBELKKINYQVDSWEEMLNKAEVCHGYMDRPLNPDAPDC 304  
Qy 528 -----APL-----TFKQGT-----ALALSCMA 544  
Db 305 PATAPNKNSTKPLDMALVNLGCGCHLSRKYMHWQEELIVGGTVKNSTGKLVSAHALQTMF 364  
Qy 545 DYCAPVFPFLAIGGYKGDYSEAEALIMTFLSNYPAGDPRLAQAKLWEAEFLMEWAFQ 604  
Db 365 QLMTPKQMY---BHFQGYEY-----VSHINWNE---DKAAAILAEAWQTYVEVHQS 411  
Qy 605 RRMAGNFQVFTAEERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSRSRVNVD 664  
Db 412 AQNSTQKVSFTT--TILDDILKSFSDSVIRVAGVLLMLAYACLTM--LRWD--CSKS 465  
Qy 665 KATLGLGVAVVLGAVMAAMGFFSYLGISSVLQVPPFLVLSVAGNADNIFIVLEYQRL 724  
Db 466 QGAVLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGVDDVFLLAHAFSET 525  
Qy 725 PRPGPREVHIGALGRVAPSMLLCSLSEAI CFELGALTMPBVRTEALTSLGLVILDF 784  
Db 526 GQNKRIPFEDRTGECUKRTGASVALTSISNVTAFFMAALIPILPALRAFSLOAAVVVVFNE 585  
Qy 785 LLOMSAFVALLSDSKQEAASRLDVCCC-----VKPOEL-----PPPG 822  
Db 586 ANVLLFPALLSMDLTYREDRRLDIFCCFTSPCVSRVIOVEPOAYTDTHTNTRYSPPPY 645  
Qy 823 QGBGL----- 827  
Db 646 SSHFAHETQITMSTQVLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSQSPST 705  
Qy 828 -----LLGFFQKAYAPFLHWHITRGVVLVLLFLALFGVSL 861  
Db 706 SSTRDLSQFSDSLHCLPCTKTWLTSSFAEKHYAPFLKPKAKVVVIFLFLGLLGVSL 765  
Qy 862 YSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTGLGNFSSEAGMNAICS 921  
Db 766 YGTRVRDGLDLDIVPRETREYDFIAAQFYFSF-----YNN----- 803  
Qy 922 SAGCNFSFTQKIQYAT-----EFPEQSYLAIPASS-----WVDDFDWL----- 961  
Db 804 -----YIVTQKADYPMIQLHLLYDLHRSFNSVNVVLEENKQLPKWMLHYFRDMLQGLQD 857  
Qy 962 -----TPSS-----CCRLYISGNKDKFCPSTVNSLNCNKCMSITWGS 1000  
Db 858 AFSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVDADGI 913  
Qy 1001 VRPSVEQPHKYL--PWFLND-----RPNIKCPKGGGLAAY--STSNNLTSQDQ 1043  
Db 914 INPSA--FYIYLTAVSNDFPVAYAASQANIRPHRPWVHDK--ADYMPETRLRIPAE 968  
Qy 1044 VLASRFWAYHKPLKNSQDYTEALRAARELANITA--DLRKVPGTDPAPFVPPYITITNVFY 1102  
Db 969 IEYAQFPFYLNGRLDTSDFEAEIKVRTICSNYTSGLSSYPNG-----YPP-----LFW 1018  
Qy 1103 EOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYMLVDTVGFMAWDIS 1162  
Db 1019 EOYIGLRHLLLFISVVVLAFTFLVCAVFLNPNWTAGII--VMVLAMTVLFGMGLIGIK 1077  
Qy 1163 YNAVSLINLVSAGVMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVFAGVAMTN 1220

1078 LSAPVPIIASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133  
1221 LPIGLVLGAKAQLIQIFPRRLMLLITLGLLHGLVPLPVLISYVG--PDVNPALALEQ- 1277  
1134 LLGVLMLAGSEPDFIVRYFAVLAAILITILGVNLGLVLLPVLISFFGYPFVPSANGNRL 1193  
1278 --KRAEEAANVAVASCPNHPHSRVSTADNIYVNHSPFGSIKG 1317  
1194 PTPSPPPPSVVRFAFAPPGHGH--SGSDSDSEYSQITVSG 1233

RESULT 14

US-09-754-032-19  
Sequence 19, Application US/09754032  
Patent No. 6921646  
GENERAL INFORMATION:  
APPLICANT: SCOTT, MATHEW P  
          GOODRICH, LISA V  
          JOHNSON, RONALD L  
TITLE OF INVENTION: Patched Genes and their Use  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/754,032  
APPLICATION NUMBER: US/09/754,032  
FILING DATE: 03-Jan-2001  
CLASSIFICATION DATA:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/540,406  
FILING DATE: 06-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: a60190-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1447 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-754-032-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;  
Best Local Similarity 22.0%; Pred. No. 1.1e-51;  
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;  
QY 225 PGOAVSGIQPLNEGVARCNESQGDVATCSQDCAASCPATARPQALDSTFVLGMPGS 284  
DB 24 PGRPAGGRRRTTGLRRRAAARDP-----YLHRSYCDAAAFALSGI---- 65  
QY 285 LVLIILICSVFAVWVILLVGFVRVAPARDKRMVDPK-----KGTSLSDKLSFSTHTLL 337  
DB 66 -----SKGKATGRKAPLWLRKAFQRLFLKLGCVIQKNC 98  
QY 338 GQFFQGWGTWASWELTILVLSVIPVVALAAGLVFTTELTDPVELWSPAPNSQARSEKAPH 397  
DB 99 GRF-----LWVGLLIFGAPVGLKAANLETNVEELWVEVGGVRVRELNYT 143

QY 398 DQHFQFPFRTN-QVILTAPNRSSYRYSLLGPKNPSGILDLDLLELLE--LQERLRLH 454  
DB 144 RQKIGBEAMFNQLMIQTPEEG-----ANVLTTEALLQHLDSALOASRVHV 190  
QY 455 QWSPSPAQRNISLQDI CYAP-----LNPDTSLYDCCINSLSLOYQNNRTLLLLTA 505  
DB 191 YMYN-----ROWKLEHLCKYSGELITETGYMDQIIEYLYPCLIITPLDCCMEGAKLSGTA 246  
QY 506 NOTLMGO-----TSQVDWKDHFY-----CAN----- 527  
DB 247 --YLLGKPLRWTFNDFLEBLELKINYOVSWEEMLNKAEVGHGYMDRPPCLNPADPDC 304  
QY 528 -----APL-----TFKQGT-----ALALSCMA 544  
DB 305 PATAPNKNSTKPLDMALVNLGGCHGLSRKYMHEQELIVGGTVKNSTGKLSVAHALQTMF 364  
QY 545 DYGA PVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWAEFLERAFQ 604  
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINNE---DKAAAILEAWQRTYVVEVHQSV 411  
QY 605 RRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYTSLALGSSYSSKSRVWVDS 664  
DB 412 AQNSTQKVLSFTT-TTLDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS 465  
QY 665 KATLGLGGVAVVLGAVMAAMGFFSYLIGIRSSLVILQVVPFLVLSVGADNIFVLEYQRL 724  
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGDVDFLLAHAFSET 525  
QY 725 PRPGEPREVIHGRALGRVAPSMMLCSLSEAICFFLGAITPMPAVRTPALTSGLAVILDF 784  
DB 526 GQNKRIPEDRGTGECUKRTGASVALTSISNVTAFMAALIPALRAFSLQAQVAVVVFNF 585  
QY 785 LIQMSAFVALLSLDSKROEASLDVCCC-----VKQDEL-----PPGS 822  
DB 586 AMVLLIFPAILSMWLYRREDRLDIFCCFTSPCVSRVQVPEQAYTDTHTNTRYSPPPY 645  
QY 823 QGEGE----- 827  
DB 646 SSHSFAHETQITMQSTVQLRTEYDPTHVYVYTTASPRSEISVQPVTVTQDTLSCQSPST 705  
QY 828 -----LLGFPKQAVAPFLHWHITRGVLLLFLALFGVSL 861  
DB 706 SSTRDLLSQFSDSSLHCLLEPPCTKWTLSFAEKHYAPFLKPKAKVWVFLFLGLIGVSL 765  
QY 862 YSMCHISVGLQDELALPKDSYLLDYFLNRYFVEGAPVYFVTILGYNFSSEAGNNAICS 921  
DB 766 YGTTVRDGLDLDIVPRETREYDFIAAQKYFSF-----YNN----- 803  
QY 922 SAGCNFSFTQKIYAT-----BPPEQSILAIPASS-----WVDDFDIDLW----- 961  
DB 804 -----YIVTQKADYPNIQHLLYDLHRSFNSVYKVMLEENKQLPKMWLHYFRDWLQGLQD 857  
QY 962 -----TPSS-----CCRLYISGPNKDFPCPSTVNSLCKNCSITMGS 1000  
DB 858 APDSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVADAGI 913  
QY 1001 VRPSVEQTHKYL-PWFLND-----RNIKCPKGLLAAY--STSVNLTSQDQ 1043  
DB 914 INPSA--FYIYLTAVWSNDPVAYAAASQANIRHPRPEWVHDK---ADYMPETRLRIPAAEP 968  
QY 1044 VLASRFMAHYKPKNSQDYTEALRAARELANITA-DLRKVPGTDPAPAEVFPYITTVFY 1102  
DB 969 IEYAGFPYINGLRTSDTVEAEIKVYTCISNYTSLGSSYPNG-----YPF-----LFW 1018  
QY 1103 EOYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLGLNLLSIVMLVDTVTFGMALWDIS 1162  
DB 1019 EOYIGLRHLLLFISVWLACTFLVCVAVFLNFWTAGII-VVMVLAJMTVELFCGMGLIGIK 1077  
QY 1163 YNAVSLINLVSAVGSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVPAVAMTN 1220  
DB 1078 LSAPVPIIASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133

```
Qy 1221 LPGAIVLGLAKAQLQIIFERLNLITLLGLLGLVFLPVLSYVG--PDVNPALALEQ- 1277
Db 1134 LLGVMLAGSEDFIVRVFFAVLAILTLLGLVGLVLLPVLSFFGYPVSPANGLNRL 1193
Qy 1278 --KRAEEAAVAVMASCNHPNPSRVSTADNIYVNHSEFSIKG 1317
Db 1194 PTPSPPEPPSVVRFAFMPGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 15
US-08-916-140-19
; Sequence 19, Application US/08916140
; Patent No. 6946257
; GENERAL INFORMATION:
; APPLICANT: Scott Mathew P.
; APPLICANT: Goodrich, Lisa V.
; APPLICANT: Johnson, Ronald L.
; APPLICANT: Epstein, Ervin Jr.
; TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO
; FILE REFERENCE: CIBT-P04-203
; CURRENT APPLICATION NUMBER: US/08/916,140
; CURRENT FILING DATE: 1997-08-21
; PRIOR APPLICATION NUMBER: US 08/656,055
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: US 08/540,406
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: US 08/319,745
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-08-916-140-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PQAVGSGIQIPLNEGVARCNESQGDVATCSQDCAACPAIARPOALDSTFYLGQMPGS 284
Db 24 PGRPAGGRRRTTGLRAAAPDRD-----YLHPSYCDAAFALEQI--- 65
Qy 285 LVLIILCSVFAVVTILLVGRFVAPARDKSRMVDK-----KGTSLSDKLSFSTHLL 337
Db 66 -----SKGRATGRKAPFLWLRKFPQRLFLKGCYIQKNC 98
Qy 338 GQFFQGWGTWASWPLTILVSVIPVALAAGLVFTELTDPVLSAPNSQARSEKAFH 397
Db 99 GK-----LVGLLIFGAPVGLKKAANLETNVEELWVEVGRVSRRELNYT 143
Qy 398 DQHFQFFPRTN-QVILTAPNRSYRYSLLGLPKNFSGILDLLELLE--LQERLHL 454
Db 144 RQKIGEEAFNPQMIQTPKEG-----ANVLTEALLQHLDSALQASRVH 190
Qy 455 QWSEPAQRNLSLQDICVAP-----LNPDTNLSYDCCINSLLQYFQNNRLLTLTA 505
Db 191 YMYN---RQWKLHLCKYSGELITETGYMDQIIEYLPCLIIITPLDCFWRGAKLQSGTA 246
Qy 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFDPLEFLEELKKNYQVDSWEMLNKAEBVGHYMDRFLNAPDPC 304
Qy 528 -----APL-----TFKDG-----ALALSCMA 544
Db 305 PATAPKNKSTPLDMALVNGCGHLSRKYMHWOBEELIVGGTVKNSTGLVSAHALQTMF 364
Qy 545 DYGAFFVFLAIGGKDYSEAEALIMTFISNNYPAGDPRLAQAKLWEAFLEEMRAQ 604
Db 365 QLMTPKQMY---EHFKGYEY---VSHINWNE-----DKAAALEAWQRTYEVVHQSV 411
```

```
Qy 605 RRMAGMQVTFPTAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSYSSWRVMWDS 664
Db 412 AQNSTQKVLSEFTT--TTLDDILKGSFSDSVIRVASGYLLMLAYACLTM---LRMD--CSKS 465
Qy 665 KATLGLGGVAVVLGAVNAAMGFFSYLGIRSSLVLQVVPFVLVLSVGADNFIIFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALUGVGDDVFLLAHAFSET 525
Qy 725 PRPGEPEVHI GRALGRVAPSMLLCSLSBAICFFFLGALTMPFAVRTFALTSGIAVLIDF 784
Db 526 GQNKRIPIFEDRTGEBCKRTGASVALTSISNVTAFMAALIPALRAFSLQAAVVVVFNF 585
Qy 785 LLQMSAFVALLSLDSKQEARSLDVCCC-----VKQEL-----PPG 822
Db 586 AMVLLIPAILSMDLVYRREDRLDIFCCFTSPCVSRVIOVEPOAYTHTDNTRYSPPPY 645
Qy 823 QGEG----- 827
Db 646 SSHSFAHETQITWQSTVOLTEYDPHTHYVYVYTAEPSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFFQKAYAPFLLHMITRQVGVLLFLALFGVSL 861
Db 706 SSTRDLLSQFSDSLHCLPEPCTKWTLSSPAETHYAPFLKPKAKVWVIFLFLGLLQVSL 765
Qy 862 YSNCHISVGLDOELALPKDSYLLDYFLFNRYFEVGAQVYFVTTLGYNFSSEAGMNAICS 921
Db 766 YGTTVRVDGLDLTDIVPRETDEYDIAAQPKYFSF-----YNN----- 803
Qy 922 SAGCNFSTFKIQYAT-----EFPEQSYLAIPASS-----WVDDFIDWL----- 961
Db 804 -----YIVTQKADYFNIQHLLYDLHRSFSNKKYVMLEENKQLPKWHLHYFRDWLQGLQD 857
Qy 962 -----TPSS-----CORLYISGNKDKFCPSTVNSLNCNKMCSITMGS 1000
Db 858 AFDSDMETGKIMPNNYKNGSDGVLAYKLLVQTGRDK--PIDISQLT--KORLVDADGI 913
Qy 1001 VRPSVEOFHKYL--PWFLND-----RPNKCPKGGGLAAV--STSVMNLTSDGQ 1043
Db 914 INFSA--FYIYLTAWVNSNDPVAYAASQANTRPHRPEWHDK---ADYMPETRLRIPAE 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPGTDPAFEPVYITITNVFY 1102
Db 969 IEVAQPFYPLNGRLDTSDFVEAIEKVRTICSNVTSLSGLSSYPNG-----YPP---LFW 1018
Qy 1103 EGYLTILPEGLFMLSCLVETPAVSCLLGLDRLSGLNLLSIVMILVDTVGFMAIWDIS 1162
Db 1019 EGYIGLRHLLFLFISVVLACTFLVCVFLNPNWTAGII--VMVLALMTVELFGMMGLIGIK 1077
Qy 1163 YNAVSLINLVASGMSVEFVSHITRSP--AISTKPTWLERAKEATISMSGSAVAGVAMTN 1220
Db 1078 LSAVPVILLASVIGVEFTVHVVALAFLTAIGDKN---REAVLALEHMFAPVLGD-AVST 1133
Qy 1221 LPGILVLGLAKAQLQIIFPFRMLNLLITLLGLLHGLVFLPVLSYVG--PDVNPALALEQ- 1277
Db 1134 LLGVMLAGSEDFIVRVFFAVLAILTLLGLVGLVLLPVLSFFGYPVSPANGLNRL 1193
Qy 1278 --KRAEEAAVAVMASCNHPNPSRVSTADNIYVNHSEFSIKG 1317
Db 1194 PTPSPPEPPSVVRFAFMPGHTH--SGSDSDSEYSSQTTVSG 1233
```

Search completed: April 7, 2006, 19:23:05  
Job time : 63 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:32:55 ; Search time 188 Seconds  
(without alignments)  
2960.365 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLRLAQ.....GSIKAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications\_AA\_Main:\*
- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:\*
  - 3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:\*
  - 4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
  - 6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	4	US-10-621-758A-4
2	6909	100.0	1332	4	US-10-663-208A-4
3	6909	100.0	1332	4	US-10-646-301A-4
4	6909	100.0	1332	4	US-10-736-769-4
5	6896	99.8	1332	4	US-10-239-316-9
6	6872.5	99.5	1359	4	US-10-621-758A-44
7	6872.5	99.5	1359	4	US-10-663-208A-44
8	6872.5	99.5	1359	4	US-10-646-301A-44
9	6872.5	99.5	1359	4	US-10-736-769-44
10	6536	94.6	1344	5	US-10-450-763-53052
11	5421.5	78.5	1331	4	US-10-621-758A-2
12	5421.5	78.5	1331	4	US-10-663-208A-2
13	5421.5	78.5	1331	4	US-10-646-301A-2
14	5421.5	78.5	1331	4	US-10-736-769-2
15	5407	78.3	1333	4	US-10-621-758A-12
16	5407	78.3	1333	4	US-10-663-208A-12
17	5407	78.3	1333	4	US-10-646-301A-12
18	5407	78.3	1333	4	US-10-736-769-12
19	4466	64.6	982	5	US-10-450-763-53050
20	2402.5	34.8	1278	4	US-10-208-731-2
21	2402.5	34.8	1278	4	US-10-741-601-530
22	2402.5	34.8	1278	5	US-10-741-600-1542
23	2402.5	34.8	1278	5	US-10-756-149-4924
24	2385	34.5	1319	4	US-10-208-731-4
25	2262	32.7	1287	6	US-11-097-143-12003
26	1864.5	27.0	1223	6	US-11-097-143-2679
27	1329	19.2	1170	4	US-10-208-731-6

28	1078	15.6	1296	4	US-10-208-731-9	Sequence 9, Appli
29	942	13.6	194	4	US-10-239-316-8	Sequence 8, Appli
30	900.5	13.0	492	4	US-10-424-599-211862	Sequence 211862,
31	859.5	12.4	541	4	US-10-424-599-189288	Sequence 189288,
32	651.5	9.4	1447	2	US-08-954-701A-19	Sequence 19, Appli
33	651.5	9.4	1447	3	US-09-898-533-5	Sequence 5, Appli
34	651.5	9.4	1447	3	US-09-754-032-19	Sequence 19, Appli
35	651.5	9.4	1447	4	US-10-421-446-19	Sequence 19, Appli
36	651.5	9.4	1447	4	US-10-791-844-6	Sequence 6, Appli
37	646.5	9.4	1434	2	US-08-954-701A-10	Sequence 10, Appli
38	646.5	9.4	1434	3	US-09-754-032-10	Sequence 10, Appli
39	646.5	9.4	1434	4	US-10-421-446-10	Sequence 10, Appli
40	628	9.1	1296	4	US-10-302-279-60	Sequence 60, Appli
41	598.5	8.7	933	4	US-10-415-934-3	Sequence 3, Appli
42	597	8.6	821	5	US-10-890-776A-4805	Sequence 4805, Ap
43	588	8.5	954	4	US-10-060-756A-3	Sequence 3, Appli
44	588	8.5	954	5	US-10-890-776A-3	Sequence 3, Appli
45	584	8.5	1203	3	US-09-990-046-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-621-758A-4  
; Sequence 4, Application US/10621758A  
; Publication No. US20040093629A1  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JBO1603-K-US  
; CURRENT APPLICATION NUMBER: US/10/621,758A  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-621-758A-4

Query Match	100.0%;	Score 6909;	DB 4;	Length 1332;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1332;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAEAGLRGWLWALLRLRLAQSEPYTTIHQGYCAFYDECGKPELSGSLMTLSNVSCLSN	60	
Db	1	MAEAGLRGWLWALLRLRLAQSEPYTTIHQGYCAFYDECGKPELSGSLMTLSNVSCLSN	60	
Qy	61	TPARKITGDHLLILQKICPRLYTGPTNTQACCSAKOLVLEASLSITKALLTRCPACSDNF	120	
Db	61	TPARKITGDHLLILQKICPRLYTGPTNTQACCSAKOLVLEASLSITKALLTRCPACSDNF	120	
Qy	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVAIYAFYQHSFPAEQSYDSCSRVRVPAA	180	
Db	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVAIYAFYQHSFPAEQSYDSCSRVRVPAA	180	
Qy	181	ATLAVGTGCVGVGSCALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240	
Db	181	ATLAVGTGCVGVGSCALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240	
Qy	241	ARCNSQGGDDVATCSQDCAASCPAARQALDSTFFYLGQMPGSLVLIILCSFVAVVTI	300	
Db	241	ARCNSQGGDDVATCSQDCAASCPAARQALDSTFFYLGQMPGSLVLIILCSFVAVVTI	300	
Qy	301	LLVGRVPAPARDKSNVDPKKGTSLSDKLSFSTHLLGQFQCGWGTWASWPLTILVLSV	360	
Db	301	LLVGRVPAPARDKSNVDPKKGTSLSDKLSFSTHLLGQFQCGWGTWASWPLTILVLSV	360	

```
QY 361 IPVVALAAGLVFTELTTPVELWSAPNSOARSEKAFHDHFGPFRFTNQVILTAPNRSSY 420
DB 361 IPVVALAAGLVFTELTTPVELWSAPNSOARSEKAFHDHFGPFRFTNQVILTAPNRSSY 420
QY 421 RYDSLGLGKPNFSGILDLLELLELOERLRLHLOVSPQAORNIISLODICYAPLNPNT 480
DB 421 RYDSLGLGKPNFSGILDLLELLELOERLRLHLOVSPQAORNIISLODICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVWDKDHFLYCANAPLTFKDGTAAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVWDKDHFLYCANAPLTFKDGTAAL 540
QY 541 SCMDYGAQVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFLIEM 600
DB 541 SCMDYGAQVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFLIEM 600
QY 601 RAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660
DB 601 RAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660
QY 661 MVDSKATILGLGVAIVLGAVMAMGFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MVDSKATILGLGVAIVLGAVMAMGFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSAEICFFLIGALTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSAEICFFLIGALTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGELLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGELLGFFQKAYAPFL 840
QY 841 LHMTRGVVLLFALFGVLSYSMCHISVGLDQELALPKDSYILDYFLFNRYFEVGA 900
DB 841 LHMTRGVVLLFALFGVLSYSMCHISVGLDQELALPKDSYILDYFLFNRYFEVGA 900
QY 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASSWDDFDW 960
DB 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASSWDDFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKMSTMGSVRPSVEQPHKYLPLFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKMSTMGSVRPSVEQPHKYLPLFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLASFMAHYHPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLASFMAHYHPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGFMAWDIISYNAVSLINLUSVAGMSVEPVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIVMLVDTVGFMAWDIISYNAVSLINLUSVAGMSVEPVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQILQIIFFFELNLLITLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQILQIIFFFELNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPNPALAEQKRAEEAAVAVMVVASCNPHSPRSVSTADNIYVNHSPFEGSIKAGA 1320
DB 1261 ILSYVGPDPNPALAEQKRAEEAAVAVMVVASCNPHSPRSVSTADNIYVNHSPFEGSIKAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332
```

RESULT 2

US-10-663-208A-4

; Sequence 4, Application US/10663208A

```
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J0101603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-663-208A-4
```

```
Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MABAGLGLWLLMALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MABAGLGLWLLMALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLIILQKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLIILQKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNOSLFINTVRVAQLGAGOLPAVVAYEAFYQHSFAEQSDYSCSRVRPAA 180
DB 121 VNLHCHNTCSPNOSLFINTVRVAQLGAGOLPAVVAYEAFYQHSFAEQSDYSCSRVRPAA 180
QY 181 ATLAAGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGIQLNEGV 240
DB 181 ATLAAGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGIQLNEGV 240
QY 241 ARCNESQGDVATCQDCQCAASCPATARPQALDSTFVLGMPGSLVLIILCSVEAVVTI 300
DB 241 ARCNESQGDVATCQDCQCAASCPATARPQALDSTFVLGMPGSLVLIILCSVEAVVTI 300
QY 301 LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPVELWSAPNSOARSEKAFHDHFGPFRFTNQVILTAPNRSSY 420
DB 361 IPVVALAAGLVFTELTTPVELWSAPNSOARSEKAFHDHFGPFRFTNQVILTAPNRSSY 420
QY 421 RYDSLGLGKPNFSGILDLLELLELOERLRLHLOVSPQAORNIISLODICYAPLNPNT 480
DB 421 RYDSLGLGKPNFSGILDLLELLELOERLRLHLOVSPQAORNIISLODICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVWDKDHFLYCANAPLTFKDGTAAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVWDKDHFLYCANAPLTFKDGTAAL 540
QY 541 SCMDYGAQVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFLIEM 600
DB 541 SCMDYGAQVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFLIEM 600
QY 601 RAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660
DB 601 RAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660
QY 661 MVDSKATILGLGVAIVLGAVMAMGFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
```

Db 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLQVVPFLVSVGADNIFIVLE 720  
QY 721 YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAIICFFLGALTMPAVRTFALTSGLAV 780  
Db 721 YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAIICFFLGALTMPAVRTFALTSGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840  
QY 841 LHWITRGVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900  
Db 841 LHWITRGVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900  
QY 901 YFVTTILGYNFSSAGNNAICSSAGCNFFSTQKIQYATFPPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNFSSAGNNAICSSAGCNFFSTQKIQYATFPPEQSYLAIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKYLPMFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKYLPMFLNDRP 1020  
QY 1021 NIKCPKGLAAYSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080  
Db 1021 NIKCPKGLAAYSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080  
QY 1081 RKVPGTDPAPFVFPYITIVVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
Db 1081 RKVPGTDPAPFVFPYITIVVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
QY 1141 NLLSIVMILVDTGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200  
Db 1141 NLLSIVMILVDTGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200  
QY 1201 AKEATISMSAVFAGVAMTNLPGILVLGLAKAQLIQIIPFRLNLLITLGLHLGLVFLPV 1260  
Db 1201 AKEATISMSAVFAGVAMTNLPGILVLGLAKAQLIQIIPFRLNLLITLGLHLGLVFLPV 1260  
QY 1261 ILSYVGPDVNPALAEQKRAEAAVAVMVASCPNHPRSRVSTADNIVNHSFSGSKGAGA 1320  
Db 1261 ILSYVGPDVNPALAEQKRAEAAVAVMVASCPNHPRSRVSTADNIVNHSFSGSKGAGA 1320  
QY 1321 ISNFLPNNGRQF 1332  
Db 1321 ISNFLPNNGRQF 1332

RESULT 3  
US-10-646-301A-4  
; Sequence 4, Application US/10646301A  
; Publication No. US20040137467A1  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JB01603-K1-US  
; CURRENT APPLICATION NUMBER: US/10/646,301A  
; CURRENT FILING DATE: 2003-08-22  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 10/621,758  
; PRIOR FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-646-301A-4

Query Match 100.0%; Score 6909; DB 4; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAEAGLRGWLWALLRLRLAQSEBPTTIHQPGYCAFYDECGKPELSSGSLMTLSNVCSLSN 60  
Db 1 MAEAGLRGWLWALLRLRLAQSEBPTTIHQPGYCAFYDECGKPELSSGSLMTLSNVCSLSN 60  
QY 61 TPARKITGDHLLILLOKICPRLTYGTNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILLOKICPRLTYGTNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120  
QY 121 VNLHCHTCSPNQSLFPIVTRVAQLGAGOLPAVYAEFYQHSFABQSYDSCSRVRPAA 180  
Db 121 VNLHCHTCSPNQSLFPIVTRVAQLGAGOLPAVYAEFYQHSFABQSYDSCSRVRPAA 180  
QY 181 ATLAGVTMGVYGSALCNAQRWLNFGDTGNGLAPLDTIFHLLERQOAVSGIQLNEGV 240  
Db 181 ATLAGVTMGVYGSALCNAQRWLNFGDTGNGLAPLDTIFHLLERQOAVSGIQLNEGV 240  
QY 241 ARCNEISQGGDDVATCSQDCAASCPAIARPOALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300  
Db 241 ARCNEISQGGDDVATCSQDCAASCPAIARPOALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300  
QY 301 LLVGRFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360  
Db 301 LLVGRFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360  
QY 361 IPVVALAAGLVTELTTPVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTAPNRSSY 420  
Db 361 IPVVALAAGLVTELTTPVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTAPNRSSY 420  
QY 421 RYDSSLGPKNPSGILDLDLLELELOERLRLHLOWSPQAORNTSLQDICYAPLNPDNT 480  
Db 421 RYDSSLGPKNPSGILDLDLLELELOERLRLHLOWSPQAORNTSLQDICYAPLNPDNT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTFKDGTTALAL 540  
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTFKDGTTALAL 540  
QY 541 SCMDYGPAPVFPFLAIGYKGYSEAEALIMTFSLNYPAGDPRLAOKLWEEAPLEEM 600  
Db 541 SCMDYGPAPVFPFLAIGYKGYSEAEALIMTFSLNYPAGDPRLAOKLWEEAPLEEM 600  
QY 601 RAFORMAGMFOVTTFAERSLEDEINRTTAEDLPITFATSYIVIFLYISLALGSYSWSRV 660  
Db 601 RAFORMAGMFOVTTFAERSLEDEINRTTAEDLPITFATSYIVIFLYISLALGSYSWSRV 660  
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLQVVPFLVSVGADNIFIVLE 720  
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLQVVPFLVSVGADNIFIVLE 720  
QY 721 YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAIICFFLGALTMPAVRTFALTSGLAV 780  
Db 721 YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAIICFFLGALTMPAVRTFALTSGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840  
QY 841 LHWITRGVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900  
Db 841 LHWITRGVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900  
QY 901 YFVTTILGYNFSSAGNNAICSSAGCNFFSTQKIQYATFPPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNFSSAGNNAICSSAGCNFFSTQKIQYATFPPEQSYLAIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKYLPMFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKYLPMFLNDRP 1020  
QY 1021 NIKCPKGLAAYSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080



Db	1321	ISNFPNNGRQF	1332		
RESULT 5					
US-10-239-316-9					
; Sequence 9, Application US/10239316					
; Publication No. US20030125253A1					
; GENERAL INFORMATION:					
; APPLICANT: TANIYAMA, Yoshio					
; APPLICANT: KITA, Shunbun					
; APPLICANT: SATOMI, Tomoko Komiya					
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And					
; FILE REFERENCE: 2703USOP					
; CURRENT APPLICATION NUMBER: US/10/239,316					
; CURRENT FILING DATE: 2002-09-19					
; PRIOR APPLICATION NUMBER: PCT/JP01/02279					
; PRIOR FILING DATE: 2001-03-22					
; PRIOR APPLICATION NUMBER: JP2000-088595					
; PRIOR FILING DATE: 2000-03-24					
; NUMBER OF SEQ ID NOS: 59					
; SEQ ID NO 9					
; LENGTH: 1332					
; TYPE: PRT					
; ORGANISM: Human					
US-10-239-316-9					
Query Match					
Best Local Similarity 99.8%; Score 6896; DB 4; Length 1332;					
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN	60		
Db	1	MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN	60		
QY	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDNF	120		
QY	121	VNLHCNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFABQSDYSCSRVRVPA	180		
Db	121	VNLHCNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFABQSDYSCSRVRVPA	180		
QY	181	ATLAVGTGCVGYGSALCNAQRWLNFGQDTGNGLAPLDTIFHLLRPGQAVGSGIQPLNEGV	240		
Db	181	ATLAVGTGCVGYGSALCNAQRWLNFGQDTGNGLAPLDTIFHLLRPGQAVGSGIQPLNEGV	240		
QY	241	ARCNESQGDVATCSQDCAASCPIARPOALDSTFYLGMPGSLVLIILCSVPVAVTI	300		
Db	241	ARCNESQGDVATCSQDCAASCPIARPOALDSTFYLGMPGSLVLIILCSVPVAVTI	300		
QY	301	LLVGRVPAPARDKSRWDPKGTSLSDKLSSTHTLLGQFGQWGTWVWASWPLTLLVLSV	360		
Db	301	LLVGRVPAPARDKSRWDPKGTSLSDKLSSTHTLLGQFGQWGTWVWASWPLTLLVLSV	360		
QY	361	IPVVALAAGLVTELTTPVELWSAPNSQARSEKAFHQHGFPPRTNQVILTAPNRSSY	420		
Db	361	IPVVALAAGLVTELTTPVELWSAPNSQARSEKAFHQHGFPPRTNQVILTAPNRSSY	420		
QY	421	RYDSLGLGKPNFSGILDLDLLELELERLHLQVWSPQARNISLQDICVAPLNPNPT	480		
Db	421	RYDSLGLGKPNFSGILDLDLLELELERLHLQVWSPQARNISLQDICVAPLNPNPT	480		
QY	481	SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKDHPLYCANAPLTKDGTALAL	540		
Db	481	SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKDHPLYCANAPLTKDGTALAL	540		
QY	541	SCMADYGAPVFPFLAIGYKGDYGEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM	600		
Db	541	SCMADYGAPVFPFLAIGYKGDYGEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM	600		
QY	601	RAFQRMAGMFOVTTAERSLEDEINRTTAEPLPATSIVIFLYISLALGSYSSWSRV	660		
Db	601	RAFQRMAGMFOVTTAERSLEDEINRTTAEPLPATSIVIFLYISLALGSYSSWSRV	660		

Db	601	RAFQRMAGMFOVTTAERSLEDEINRTTAEPLPATSIVIFLYISLALGSYSSWSRV	660	
QY	661	MYDSKATILGLGVAVVILGAVMAAMGFPSYILGRSSIVILQVVPELVLSVGADNIFIVLE	720	
Db	661	MYDSKATILGLGVAVVILGAVMAAMGFPSYILGRSSIVILQVVPELVLSVGADNIFIVLE	720	
QY	721	YQRLPRRPGEPREVIHGRALGRVAFSMLLCSLSEACFFELGALTMPAVRTFALTSLGLAV	780	
Db	721	YQRLPRRPGEPREVIHGRALGRVAFSMLLCSLSEACFFELGALTMPAVRTFALTSLGLAV	780	
QY	781	ILDFFLQNSAFVALLSLDSKROEASRLDVCCCVKQOELPPQCGEGLLIGFQKAYAPFL	840	
Db	781	ILDFFLQNSAFVALLSLDSKROEASRLDVCCCVKQOELPPQCGEGLLIGFQKAYAPFL	840	
QY	841	LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVAPV	900	
Db	841	LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVAPV	900	
QY	901	YFVTTLGYNFSSEAGNAICSSAGCNFSFTQIQYATEFFPEQSYLAIPASSWVDDFDM	960	
Db	901	YFVTTLGYNFSSEAGNAICSSAGCNFSFTQIQYATEFFPEQSYLAIPASSWVDDFDM	960	
QY	961	LTPSSCCRLYISGPNKXFCPSTVNSLNLCKNCHSITWGSVRPSVEQPHKYLFWFLNDRP	1020	
Db	961	LTPSSCCRLYISGPNKXFCPSTVNSLNLCKNCHSITWGSVRPSVEQPHKYLFWFLNDRP	1020	
QY	1021	NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL	1080	
Db	1021	NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL	1080	
QY	1081	RKVPCTDPAFEPVPTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140	
Db	1081	RKVPCTDPAFEPVPTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140	
QY	1141	NLLSVTMLTVDTVGNALWDISYNASVLINLVSAGMSVEFVSHITRSPAIKTKPTWLER	1200	
Db	1141	NLLSVTMLTVDTVGNALWDISYNASVLINLVSAGMSVEFVSHITRSPAIKTKPTWLER	1200	
QY	1201	AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRNLNLIITLLGLHGLVFLGPV	1260	
Db	1201	AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRNLNLIITLLGLHGLVFLGPV	1260	
QY	1261	ILSVYGPDPNPALALEQKRAEEAAVAVMVASCPNHPSPVSTADNIYVNHSPSGSIKAGA	1320	
Db	1261	ILSVYGPDPNPALALEQKRAEEAAVAVMVASCPNHPSPVSTADNIYVNHSPSGSIKAGA	1320	
QY	1321	ISNFPNNGRQF	1332	
Db	1321	ISNFPNNGRQF	1332	

RESULT 6  
US-10-621-758A-44  
; Sequence 44, Application US/10621758A  
; Publication No. US20040093629A1  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC111  
; FILE REFERENCE: JB01603-K-US  
; CURRENT APPLICATION NUMBER: US/10/621,758A  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 44  
; LENGTH: 1359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-621-758A-44

Query Match		99.5%;	Score 6872.5;	DB 4;	Length 1359;
Best Local Similarity		97.9%;	Pred. No. 0;		
Matches 1330;		Conservative	0;	Mismatches	2; Indels 27; Gaps 1;
Qy	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECCKNPGLSGMLTSLNVSCLSN	60		
Db	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECCKNPGLSGMLTSLNVSCLSN	60		
Qy	61	TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA	180		
Db	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA	180		
Qy	181	ATLAVGTMCGVYSGALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV	240		
Db	181	ATLAVGTMCGVYSGALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV	240		
Qy	961	LTPSSCCRLYISGPNKDKFCSTVNSLNCNCHSITMGSVRPSVEQPHKYLPHFLNDRP	1020		
Db	961	LTPSSCCRLYISGPNKDKFCSTVNSLNCNCHSITMGSVRPSVEQPHKYLPHFLNDRP	1020		
RESULT 7					
US-10-663-208A-44					
; Sequence 44, Application US/10663208A					
; Publication No. US20040132058A1					
; GENERAL INFORMATION:					
; APPLICANT: Altmann, Scott W					
; APPLICANT: Wang, Luquan					
; APPLICANT: Graziano, Michael					
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF					
; FILE REFERENCE: JB01603K2 US					
; CURRENT APPLICATION NUMBER: US/10/663, 208A					
; CURRENT FILING DATE: 2003-09-16					
; PRIOR APPLICATION NUMBER: 60/397,442					
; PRIOR FILING DATE: 2002-07-19					
; PRIOR APPLICATION NUMBER: 10/621,758					
; PRIOR FILING DATE: 2003-07-17					
; PRIOR APPLICATION NUMBER: 10/646,301					
; NUMBER OF SEQ ID NOS: 50					
; SOFTWARE: PatentIn version 3.1					
; SEQ ID NO 44					
; LENGTH: 1359					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-663-208A-44					
Query Match		99.5%;	Score 6872.5;	DB 4;	Length 1359;
Best Local Similarity		97.9%;	Pred. No. 0;		
Matches 1330;		Conservative	0;	Mismatches	2; Indels 27; Gaps 1;
Qy	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECCKNPGLSGMLTSLNVSCLSN	60		
Db	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECCKNPGLSGMLTSLNVSCLSN	60		
Qy	61	TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA	180		
Db	121	VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA	180		
Qy	181	ATLAVGTMCGVYSGALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV	240		
Db	181	ATLAVGTMCGVYSGALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV	240		
Qy	241	ARCNESQDDVATCSQDCAACCPAIAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300		
Db	241	ARCNESQDDVATCSQDCAACCPAIAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300		
Qy	301	LLVGRVAPARDKSKMDPKGTSLSDKLSFTSTHLLGQFQGWGTWVASHPILILVSV	360		
Db	301	LLVGRVAPARDKSKMDPKGTSLSDKLSFTSTHLLGQFQGWGTWVASHPILILVSV	360		
Qy	361	IPVVALAAGLVFTELTDTPVELWAPNSQARSEKAFHDQHGFPFRTNOVILTAPNRSSY	420		
Db	361	IPVVALAAGLVFTELTDTPVELWAPNSQARSEKAFHDQHGFPFRTNOVILTAPNRSSY	420		
Qy	421	RYDSLLLGKPNFSGILDLDLLELLEQLERLHQLVMSPEAQRNISLQDICYAPLNPDNT	480		
Db	421	RYDSLLLGKPNFSGILDLDLLELLEQLERLHQLVMSPEAQRNISLQDICYAPLNPDNT	480		
Qy	481	SLYDCCINSILQYFQNNRTALLTANQTLMGQTSQVDKQDFLYCANAPLTFKQGTALAL	540		
Db	481	SLYDCCINSILQYFQNNRTALLTANQTLMGQTSQVDKQDFLYCANAPLTFKQGTALAL	540		
Qy	541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTSLNYPAGDPRLAQAKLWEAFLEEM	600		
Db	541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTSLNYPAGDPRLAQAKLWEAFLEEM	600		
Qy	601	RAFQRMAGMFQVTFMAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGSYSWSRV	660		
Db	601	RAFQRMAGMFQVTFMAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGSYSWSRV	660		
Qy	661	MVDSKATILGCGVAVVLGAVNAAAGFFSYLGRSLVILQVVPFLVLSVGDNIFIFVLE	720		
Db	661	MVDSKATILGCGVAVVLGAVNAAAGFFSYLGRSLVILQVVPFLVLSVGDNIFIFVLE	720		
Qy	721	YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSAICFFFLGALTPMPAVRTFALTSLGLAV	780		
Db	721	YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSAICFFFLGALTPMPAVRTFALTSLGLAV	780		
Qy	781	ILDFFLQMSAFVALLSLDSKQKQASRLDVCCCKPQELPPPGQEGLLLGFFQKAYAPFL	840		
Db	781	ILDFFLQMSAFVALLSLDSKQKQASRLDVCCCKPQELPPPGQEGLLLGFFQKAYAPFL	840		
Qy	841	LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLDYFLFLNRYFEVGA	900		
Db	841	LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLDYFLFLNRYFEVGA	900		
Qy	901	YFVTTLGYNFSSAAGMNAICSSAGCNFPSTQKIQYATEFPEQSYLAIPASSWDDFDW	960		
Db	901	YFVTTLGYNFSSAAGMNAICSSAGCNFPSTQKIQYATEFPEQSYLAIPASSWDDFDW	960		
Qy	961	LTPSSCCRLYISGPNKDKFCSTVNSLNCNCHSITMGSVRPSVEQPHKYLPHFLNDRP	1020		
Db	961	LTPSSCCRLYISGPNKDKFCSTVNSLNCNCHSITMGSVRPSVEQPHKYLPHFLNDRP	1020		

Db	241	ARCNESQGDVATCSCQDCCAASCPAIAARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI	300
Qy	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQGGTGWASWPLTILVLSV	360
Db	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQGGTGWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFFRTNQVILTPAPNRSSY	420
Db	361	IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFFRTNQVILTPAPNRSSY	420
Qy	421	RYDSLILGPKNFGSILDLLELLELQERLHLQVWSPQARNISLODICYAPLNPNT	480
Db	421	RYDSLILGPKNFGSILDLLELLELQERLHLQVWSPQARNISLODICYAPLNPNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL	540
Qy	541	SCMADYGAPVPPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRPRLAQAKLWEEAFLEEM	600
Db	541	SCMADYGAPVPPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRPRLAQAKLWEEAFLEEM	600
Qy	601	RAFQRMAGMFOVTFTAERSLEDEINRTTAEDLPITATSYIVIFLYISLALGYSYSSRV	660
Db	601	RAFQRMAGMFOVTFTAERSLEDEINRTTAEDLPITATSYIVIFLYISLALGYSYSSRV	660
Qy	661	MVDSKATLGLGGVAVVLGVAWMAAGPFSYLGIRSSILVILQVVPFLVLSVGCADNIFLVL	720
Db	661	MVDSKATLGLGGVAVVLGVAWMAAGPFSYLGIRSSILVILQVVPFLVLSVGCADNIFLVL	720
Qy	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAI CFFLGAITPMPAVRTFALTSGLAV	780
Db	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAI CFFLGAITPMPAVRTFALTSGLAV	780
Qy	781	ILDFLQMSAFVALLSDSKRQSEASRLDVCCCKPQELPPPGQEGILLGFFOKAYAPFL	840
Db	781	ILDFLQMSAFVALLSDSKRQSEASRLDVCCCKPQELPPPGQEGILLGFFOKAYAPFL	840
Qy	841	LHWITRGVLLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Qy	901	YFVTTLGYNFSSEAGNNAICSSAGCNNSFTQKIQVATEFPEQSYLAIIPASSNVDDFIDW	960
Db	901	YFVTTLGYNFSSEAGNNAICSSAGCNNSFTQKIQVATEFPEQSYLAIIPASSNVDDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMSITMGSVRPSVQFHXYLPWFNDPR	1020
Db	961	LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMSITMGSVRPSVQFHXYLPWFNDPR	1020
Qy	1021	NIKCPKGLAAYSTSVNLTSDGOVL-----ASRFMAVH 1053	
Db	1021	NIKCPKGLAAYSTSVNLTSDGOVLDTVALSPRLEYSGTISAHNCNLYLLDSASRFMAVH 1080	
Qy	1054	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPFPTITNVFYEQYLTILPEGL 1113	
Db	1081	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPFPTITNVFYEQYLTILPEGL 1140	
Qy	1114	FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGFMAWMDISYNVAVSLINLVS 1173	
Db	1141	FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGFMAWMDISYNVAVSLINLVS 1200	
Qy	1174	AVGMSVEFYSHITRSPAITKPTWLERAKEATISMGSAVFAVAMTNLPGLVLGLAKAQ 1233	
Db	1201	AVGMSVEFYSHITRSPAITKPTWLERAKEATISMGSAVFAVAMTNLPGLVLGLAKAQ 1260	
Qy	1234	LIQIFPFRNLNLTLLGLHGLVFLFVLVILSYVGPVNPALALEQKAEAEAAVAVASCP 1293	
Db	1261	LIQIFPFRNLNLTLLGLHGLVFLFVLVILSYVGPVNPALALEQKAEAEAAVAVASCP 1320	
Qy	1294	NHPSRVSTADNIYVNHSPFGSIKGAGAINFLPNNGRQF 1332	

Db	1321	NHPSRVSTADNIYVNHSPFGSIKGAGAINFLPNNGRQF 1359	
RESULT 8			
US-10-646-301A-44			
; Sequence 44, Application US/10646301A			
; Publication No. US20040137467A1			
; GENERAL INFORMATION:			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JBO1603-K1-US			
; CURRENT APPLICATION NUMBER: US/10/646,301A			
; CURRENT FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; NUMBER OF SEQ ID NOS: 50			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 44			
; LENGTH: 1359			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-646-301A-44			
Query Match 99.5%; Score 6872.5; DB 4; Length 1359;			
Best Local Similarity 97.9%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;			
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDEGCKNPGLSGSLMTLSNVCSLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDEGCKNPGLSGSLMTLSNVCSLSN	60
Qy	61	TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSNQSLEFINVTRVAQLGAGOLPAWYAEAYQHSFARQSDSCSRVRPAA	180
Db	121	VNLHCHNTCSNQSLEFINVTRVAQLGAGOLPAWYAEAYQHSFARQSDSCSRVRPAA	180
Qy	181	ATLAVGTMGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVGSGIQTNEG	240
Db	181	ATLAVGTMGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVGSGIQTNEG	240
Qy	241	ARCNESQGDVATCSCQDCCAASCPAIAARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI	300
Db	241	ARCNESQGDVATCSCQDCCAASCPAIAARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI	300
Qy	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQGGTGWASWPLTILVLSV	360
Db	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQGGTGWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFFRTNQVILTPAPNRSSY	420
Db	361	IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFFRTNQVILTPAPNRSSY	420
Qy	421	RYDSLILGPKNFGSILDLLELLELQERLHLQVWSPQARNISLODICYAPLNPNT	480
Db	421	RYDSLILGPKNFGSILDLLELLELQERLHLQVWSPQARNISLODICYAPLNPNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL	540
Qy	541	SCMADYGAPVPPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRPRLAQAKLWEEAFLEEM	600
Db	541	SCMADYGAPVPPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRPRLAQAKLWEEAFLEEM	600
Qy	601	RAFQRMAGMFOVTFTAERSLEDEINRTTAEDLPITATSYIVIFLYISLALGYSYSSRV	660

||||| 601 RAQRMAGNFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660  
QY 661 MVDKATLGLGGVAVVILGAVMAAMGFSSYLGRSSVLQVVPFLVSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVVILGAVMAAMGFSSYLGRSSVLQVVPFLVSVGADNIFIFVLE 720  
QY 721 YQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGGALTPMPAVRTFALTSLGLAV 780  
Db 721 YQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGGALTPMPAVRTFALTSLGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGQEGLLIGFFQKAYAPPL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGQEGLLIGFFQKAYAPPL 840  
QY 841 LHWITRGVVLILLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHWITRGVVLILLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTILGYNPSSSEAGNNAICSSAGCNPFSTOKI QYATEFPFQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNPSSSEAGNNAICSSAGCNPFSTOKI QYATEFPFQSYLAIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKMCSITMGSVRPSVEQPHKYLPLWFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKMCSITMGSVRPSVEQPHKYLPLWFLNDRP 1020  
QY 1021 NIKCPKGGGLAAVSTSVNLTSDGVL-----ASRFMAYH 1053  
Db 1021 NIKCPKGGGLAAVSTSVNLTSDGVLDTVAILSPRLEYSGTISAHCNVLLDSASRFMAYH 1080  
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFYEQVLTILPEGL 1113  
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFYEQVLTILPEGL 1140  
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLVDTVGMALWDSYNNAVSLINLVS 1173  
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLVDTVGMALWDSYNNAVSLINLVS 1200  
QY 1174 AVGMSVEFVSHITRFAISTKFTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITRFAISTKFTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1260  
QY 1234 LIQIIFFRNLNLLITLLGLHLGVLPVLVILSYVGPDPVPALALEQKRAEAAVAAVNVASCP 1293  
Db 1261 LIQIIFFRNLNLLITLLGLHLGVLPVLVILSYVGPDPVPALALEQKRAEAAVAAVNVASCP 1320  
QY 1294 NHPSRVSTADNIVNVHSPFGSTKKGAGAINFLPNNGROF 1332  
Db 1321 NHPSRVSTADNIVNVHSPFGSTKKGAGAINFLPNNGROF 1359

## RESULT 9

US-10-736-769-44  
; Sequence 44, Application US/10736769  
; Publication No. US20040161838A1  
; GENERAL INFORMATION:  
; APPLICANT: Altman, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JB01603-K3-US  
; CURRENT APPLICATION NUMBER: US/10/736,769  
; CURRENT FILING DATE: 2003-12-16  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 10/621,758  
; PRIOR FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 10/646,301  
; PRIOR FILING DATE: 2003-08-22  
; PRIOR APPLICATION NUMBER: 10/663,208  
; PRIOR FILING DATE: 2003-09-16

; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 44  
; LENGTH: 1359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-736-769-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTTTHQPGYCAFYDECCKNPBELSGSLMTLSNVSCLSN 60  
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTTHQPGYCAFYDECCKNPBELSGSLMTLSNVSCLSN 60  
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLSEASITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLSEASITKALLTRCPACSDNF 120  
QY 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180  
Db 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180  
QY 181 ATLAVGTMCGVGSALCNAORWLNFOGDTGNGLAPLDITPHLEPGQAVSGIOPINEGV 240  
Db 181 ATLAVGTMCGVGSALCNAORWLNFOGDTGNGLAPLDITPHLEPGQAVSGIOPINEGV 240  
QY 241 ARCNEQGDVATCSQDCAASCFAIARPOALDSTFYLQMPGSLVLIITLCSVFAVVTI 300  
Db 241 ARCNEQGDVATCSQDCAASCFAIARPOALDSTFYLQMPGSLVLIITLCSVFAVVTI 300  
QY 301 LLVGFVRAPARDSKMWDPKKGTSLSDKLSFSFTHLLGQFPQGWGTWVASWPLTILVLSV 360  
Db 301 LLVGFVRAPARDSKMWDPKKGTSLSDKLSFSFTHLLGQFPQGWGTWVASWPLTILVLSV 360  
QY 361 IPVALLAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHGFPFRFTNOVILTAPNRSSY 420  
Db 361 IPVALLAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHGFPFRFTNOVILTAPNRSSY 420  
QY 421 RYDSLLGPNFSGIILDLLELLELQERLRLHQLVMSPEAQNRISLQDICYAPLNPDNT 480  
Db 421 RYDSLLGPNFSGIILDLLELLELQERLRLHQLVMSPEAQNRISLQDICYAPLNPDNT 480  
QY 481 SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKDGTALAL 540  
Db 481 SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKDGTALAL 540  
QY 541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
Db 541 SCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
QY 601 RAFOREMAGNFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660  
Db 601 RAFOREMAGNFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660  
QY 661 MVDKATLGLGGVAVVILGAVMAAMGFSSYLGRSSVLQVVPFLVSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVVILGAVMAAMGFSSYLGRSSVLQVVPFLVSVGADNIFIFVLE 720  
QY 721 YQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGGALTPMPAVRTFALTSLGLAV 780  
Db 721 YQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGGALTPMPAVRTFALTSLGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGQEGLLIGFFQKAYAPPL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGQEGLLIGFFQKAYAPPL 840  
QY 841 LHWITRGVVLILLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHWITRGVVLILLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTILGYNPSSSEAGNNAICSSAGCNPFSTOKI QYATEFPFQSYLAIPASSWVDDFIDW 960

Db	901	YFVTLLGYNFSSSEAGNNAICSSAGCNNSFTQKIQYATFPPEQSYLAIIPASSWVDDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKWCSITMGSVRPSVEQFHXYLPWFNDLRP	1020
Db	961	LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKWCSITMGSVRPSVEQFHXYLPWFNDLRP	1020
Qy	1021	NIKCPKGGGLAAYSTSVNLSDGQVL-----ASRFMAYH	1053
Db	1021	NIKCPKGGGLAAYSTSVNLSDGQVLDTVAILSPRLSEYSGTISAHCNLYLLDSASRFMAYH	1080
Qy	1054	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPYPTITNVFYEQVLTILPEGL	1113
Db	1081	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPYPTITNVFYEQVLTILPEGL	1140
Qy	1114	FMLSCLCLVPTFAVSCILGLDLRSGLNLLSIVMILVDTVGFMAWMDISYNAVSLINLVS	1173
Db	1141	FMLSCLCLVPTFAVSCILGLDLRSGLNLLSIVMILVDTVGFMAWLGISYNAVSLINLVS	1200
Qy	1174	AVGMSVEFVSHITRSPATSTKPTWLERAKEATISMGSAVPAGVAMTNLPGLVLGLAKAQ	1233
Db	1201	AVGMSVEFVSHITRSPATSTKPTWLERAKEATISMGSAVPAGVAMTNLPGLVLGLAKAQ	1260
Qy	1234	LQIIFFRNLNLTITLGLLHGLVFLPVILSYVGPVNPALALEOKRAEAAVAVMVASCP	1293
Db	1261	LQIIFFRNLNLTITLGLLHGLVFLPVILSYVGPVNPALALEOKRAEAAVAVMVASCP	1320
Qy	1294	NHPSRVSTADNIYVNHSPFGSIKGAAGISNFIPLNNGRQF	1332
Db	1321	NHPSRVSTADNIYVNHSPFGSIKGAAGISNFIPLNNGRQF	1359

RESULT 10  
US-10-450-763-53052  
; Sequence 53052, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450, 763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 53052  
; LENGTH: 1344  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (901)..(917)  
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,  
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1344)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2

US-10-450-763-53052

Query Match	94.6%;	Score 6536;	DB 5;	Length 1344;
Best Local Similarity	96.7%;	Pred. No. 0;		
Matches 1271;	Conservative	7;	Mismatches	21;
			Indels	16;
			Gaps	4;

Qy	1	MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFFDECGKNPELSGSIWTLTNSVCSLSN	60
Db	1	MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFFDECGKNPELSGSIWTLTNSVCSLSN	60

Qy	61	TPARKITGHLILLLOKICPRLYTGNTQOACCSAKQLVLSLEASISITKALLTRCPACSDNF	120
Db	61	TPARKITGHLILLLOKICPRLYTGNTQOACCSAKQLVLSLEASISITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQAGAGOLPAVAYEAFQHSFABEQSYDSCSRVRPAA	180
Db	121	VNLHCHNTCSPNQSLFINVTRVAQAGAGOLPAVAYEAFQHSFABEQSYDSCSRVRPAA	180
Qy	181	ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLIDITFHLLPEQOAVGSGIQPLNEGV	240
Db	181	ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLIDITFHLLPEQOAVGSGIQPLNEGV	240
Qy	241	ARCNSQGGDDVATCSCQDCAACP--AIAAPQALBSTFYLGQMPGSLVLIILICSVFVAV	298
Db	241	ARCNSQGGDDVATCSCQDCAACPXPAPRPTPSTW--RQMPGSLVLIILICSVFVAV	298
Qy	299	TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFQGGTGWASWPLTILVL	358
Db	299	TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFQGGTGWASWPLTILVL	358
Qy	359	SVIPVVALAAGLVFTTELTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS	418
Db	359	SVIPVVALAAGLVFTTELTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS	418
Qy	419	STRYDSSLILGPKNFSGIILDLLELLELQERLRLQVMSPEAQRNISLQDICYAPLNPD	478
Db	419	STRYDSSLILGPKNFSGIILDLLELLELQERLRLQVMSPEAQRNISLQDICYAPLNPD	478
Qy	479	NTSLYDCCINSLLQYFQNNRTLLLTANQTLNGQTSQVDMKDFLYCANAPLTFKQGTAL	538
Db	479	NTSLYDCCINSLLQYFQNNRTLLLTANQTLNGQTSQVDMKDFLYCANAPLTFKQGTAL	538
Qy	539	ALSCMADYCAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGPRLPAQAQKLEWAEAFLE	598
Db	539	ALSCMADYCAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGPRLPAQAQKLEWAEAFLE	598
Qy	599	ENRAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSVYIVFLYISLALGSYSWS	658
Db	599	ENRAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSVYIVFLYISLALGSYSWS	658
Qy	659	RVWDSKATILGGLGVAVNLGAVMAAGPFYSYIGIRSSLVILQVVPFLVSVGADNIFIFV	718
Db	659	RVWDSKATILGGLGVAVNLGAVMAAGPFYSYIGIRSSLVILQVVPFLVSVGADNIFIFV	718
Qy	719	LEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEATCFFLGALTTPMPAVTFALTSG	778
Db	719	LEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEATCFFLGALTTPMPAVTFALTSG	778
Qy	779	AVILDFLQMSAFVALLSLDSKRQBSRLDVCCCKVQBELPPPGQEGLLGFFQKAYAP	838
Db	779	AVILDFLQMSAFVALLSLDSKRQBSRLDVCCCKVQBELPPPGQEGLLGFFQKAYAP	838
Qy	839	FLLHWITRGVV-----LILFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLNRYF	894
Db	839	FLLHWITRGVVVPSQLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLNRYF	898
Qy	895	EVGAPVYFVTTLGVNFSSEAGNNAICSSAGCNNSFTQKIQYATFPPEQSYLAIIPASSWV	954
Db	899	EVGAPVYFVTTLGVNFSSEAGNNAICSSAGCNNSFTQKIQYATFPPEQSYLAIIPASSWV	958
Qy	955	DDFIDWLTSPSSCCRLYISGPNKDKFCPSVTNSLNCNKWCSITMGSVRPSVEQFHXYLPW	1014
Db	959	DDFIDWLTSPSSCCRLYISGPNKDKFCPSVTNSLNCNKWCSITMGSVRPSVEQFHXYLPW	1018
Qy	1015	FLNDRPNIKCPKGGGLAAYSTSVNLSDGQVLASRFMAYHKLKNSQDYTEALRAARELAA	1074
Db	1019	FLNDRPNIKCPKGGGLAAYSTSVNLSDGQVLASRFMAYHKLKNSQDYTEALRAARELAA	1078
Qy	1075	NITADLRKVPGTDPAFEPYPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCILGLD	1134
Db	1079	NITADLRKVPGTDPAFEPYPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCILGLD	1138
Qy	1135	LRSGLLNLLSTVMIILVDTVGFMAWMDISYNAVSLINLVSAGVMSVEFVSHITRSPATSTK	1194

```
Db 1139 LRSGLNLNLSWILVDTGFMALWGISYNAVSLNLSVAGMSVEFVSHITRFAISTK 1198
Qy 1195 PTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLITLLGLLHG 1254
Db 1199 PTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLITLLGLLHG 1258
Qy 1255 LVFLPVILSYGPDVNPALALEOKRAEEA-----VAAMVAVSCPNHPSRVST 1301
Db 1259 LVFLPVILSYGPDVNPALALEOKRAEEGGSGHGGPLAQITPESPQLITSMST 1313

RESULT 11
US-10-621-758A-2
; Sequence 2, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2

Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGLRGWLWALLRLAQSEPTTHOPGYCAFYDECKGNKPELSGLMTLSNVCLSN 60
Db 1 MAEAWL-GWLLWALLSAAQELTPKHEAGVCTFYEECKGNPELSGLTSLSNVCLSN 59
Qy 61 TPARKITGDHLILLOKICPRLYTGPNTQ-ACCSAKQLVSLSEASITKALLTRCPACSDN 119
Db 60 TPARKITGEHALLQRICPRLYNGPNTTFACCTSKQLLSLESSITKALLTRCPACSDN 119
Qy 120 FVNILCHNTCPNOSLFNTVTRVQALGAGQLPAVVAYEAFYQHSFASQSDYSCSRVRPA 179
Db 120 FVSLHCHNTCPDQSLFNTVTRVVERGAGEPPAVVAYEAFYQSFASKAYESCQVRIPA 179
Qy 180 AATLAVGTMCGVYSALCNAORWLNFOGDTGNGLAPLDITFHLLEPGOAVSGIOPLNEG 239
Db 180 ANSLAVGSMCGVYSALCNAORWLNFOGDTGNGLAPLDITFHLLEPGOALPDGQIPLNGK 239
Qy 240 VARNESQGDVATCSCQDCQCAAFARPOALDSTFYLQMPGSLVLIILICSVFAVVT 299
Db 240 IAPCNESQGDSDAVSCQDCQCAACFPVPPPEALRPSFYMGEMPCWALIIIPAVFVLS 299
Qy 300 ILLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLS 359
Db 300 AVLRLRWVSNRNKAEQEPQEAFLPHKIKLSPTHTILGRFFQWNGTRVASWPLTILALS 359
Qy 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSKAFDQHFQFPFFRTNOVILTAPNRSS 419
Db 360 FIVVIALAAGLTFTLTDPVELWSAPKARKEKSFDEHFGFFFTNQIFVTAARNSS 419
Qy 420 YRYDSLLGPNKFGIILDLLELLELQERLHLQWSPQAQNSILODICVAPLMPDN 479
Db 420 YKYDSLLGKSNFGSILSLOFLLELLELQERLHLQWSPQAQNSILODICVAPLMPYN 479
Qy 480 TSLYDCINSILQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTSLA 539
Db 480 TSLYDCINSILQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTSLA 539
```

```
Db 480 TSLSDCCVNSLLQYFQNNRTLLMLTANQTLNGQTSVLDWKDHFLYCANAPLTKDGTSLA 539
Qy 540 LSCMADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLQAQKLWEEAFLEE 599
Db 540 LSCMADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLQAQKLWEEAFLEE 599
Qy 600 MRFAFORMAGMFOVTTFAERSLEDEINRTTAEDLPFPATSYIIVFLYISALGSYSWSR 659
Db 600 MESFORNTSDKQVAFSAERSLEDEINRTTIQDLPVFAVSYIIVFLYISALGSYSRCSR 659
Qy 660 VMYDSKATLGLGVAVVILGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFLVL 719
Db 660 VAVESKATLGLGVVILGAVLAAAGFFSYLGVPSLSLVIIQVVPFLVLAAGNIFLVL 719
Qy 720 EYORLPRRPGEPREHVHIGRALGRVAPSMILCSLSEAI CFFELGALTMPA VRTFALTSLGA 779
Db 720 EYORLPRMPGEPREAHIGRTLGSVAPSMILCSLSEAI CFFELGALTMPA VRTFALTSLGA 779
Qy 780 VILDELLQSAFVALLSLDSKROEASRLDVCCCKVQDELPPPGQEGGLILGFFQKAYAPF 839
Db 780 IILDELLQMTAFVALLSLDSKROEASRPDLVLCFSTRKPPPPKKEBGLLRFRKIIYAPF 839
Qy 840 LLEHWITRGVVLFLALFGVSLYSMSCHISVGLDQELALPKDSVLLDYFLFLNRYFEVGAP 899
Db 840 LLHRFIRPVVMLLFLTLFGANLYLMCNINVLGDELALPKDSVLLDYFLFLNRYLEVGP 899
Qy 900 VYFVTTLGVNFSSEAGNNAICSSAGCNNSFTOKIOYATEFPFQSYLAIPASSWVDDFID 959
Db 900 VYFVTTSGFNFSSEAGNNAICSSAGCKSFSLTKIOYASEFPDQSYVAIAASSWVDDFID 959
Qy 960 WLTP-SSCRLYISGPNKDKFCPSTVNSLNCNKMSTWGSVRPSVEQPHKYLPMFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDEFCESTDTSFNCLNKNWRTLGPVRTAEQPHKYLPMFLND 1019
Qy 1019 RPNKCPKGGALAAYSTSVNLTSQGVLASRFMAYHKLKNSQDYTEALRAARELAANITA 1078
Db 1020 PPNIRCPKGGALAAYSTSVNLSSDQGVIAQFMAYHKLKNSQDYTEALRAARELAANITA 1079
Qy 1079 DLKRVGTDPAFVFPFYTITNVFYEQYLIILBGLPMLSILCLVPTFAVSCILLGLDLRSG 1138
Db 1080 DLKRVGTDPNFVFPFYTITNVFYQYLYTLVPEGIPTLALCFVPTFVVCVLLGLDMCSG 1139
Qy 1139 LNLLSIVMLVDTVGFMALWDISYNAVSLINLSVAGMSVEFVSHITRFAISTKPTWL 1198
Db 1140 ILNLSIIMLVDTIGLMAVWGISYNAVSLINLTVAGMSVEFVSHITRFAISTKPTRL 1199
Qy 1199 ERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLITLLGLLHGLVFL 1258
Db 1200 ERAKDATVFMGSAVFAGVAMTNFPGLILGFAQAQLIQIFFRNLNLITLLGLLHGLVFL 1259
Qy 1259 PVTLVYGVDPNPALALEOKRAEEA VAAVMVAVSCPNHPSRVSTADNIYNHSPGSI-KG 1317
Db 1260 PVTLVYGVDPNPALVQVEEKLAEEA-AVAPEPCQPYPSPADADAN--VNYGAPBELAHG 1316
Qy 1318 AGAISNLFNNGRQF 1332
Db 1317 ANAARSSLPKSDQKF 1331
```

```
RESULT 12
US-10-663-208A-2
; Sequence 2, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
```

```
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-663-208A-2

Query Match      78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQLSEPYTTTHQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGEYTPKHEAGVCTFYEECGKPELSGLTSLNVCSLSN 59

QY 61 TPARKITGDHLLILLOKICPRLYTPNTQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
DB 60 TPARHVTGEHALLQRIQCPRLYNGPNTTFACCSKQLLSLESSMSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQAGOLPAWVAEAPYQHSFAEQSYDSCSRVRVPA 179
DB 120 FVSLCHNTCSPPQSLFINVTRVVERGAGEPPAVVAEAFYQKSPAEKAYECSQVRIIPA 179

QY 180 AATLAVGTCMGVYGSALCNAQRLNFOGDTGNGLAPLDITFHLLPEQAVGSIQPLNKG 239
DB 180 AASLAVGSMCGVYGSALCNAQRLNFOGDTGNGLAPLDITFHLLPEQAVGSIQPLNKG 239

QY 240 VARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLGOMPGSLVLIILICLSFVAVT 299
DB 240 IAPCNESQDDSAVSCQDCAASCPVIPPPEALRPSFYMGRRPGWLALIIITFAVFLVS 299

QY 300 ILLVGRVAPARDKGMVDPKXGTSLSDKLSFTHTLLGQFFQGMGTWVASWPLTILVLS 359
DB 300 AVLVRVSVNRNKNKAEGQEPAPKLPHKHKLSPTHTLGRFFQNGTRVVASWPLTILVLS 359

QY 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDHGFPPFRTNQVILTAPNRSS 419
DB 360 FIVWIALAAGLVFTLTDPVELWSAPNSQARSEKSFHDEHGFPPFRTNQVILTAPNRSS 419

QY 420 YRYSLLLPKPNFSGILDLDLLELQERLHLQVWSPAEARNISLQDICYAPLNPYN 479
DB 420 KYYSLLLPKPNFSGILSLDLLELQERLHLQVWSPAEARNISLQDICYAPLNPYN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLGQTSQVDMQKHFLYCANAPLTFKDGTLA 539
DB 480 TSLSDCCVNSLLQYFQNNRTLLMLTANQTLGQTSQVDMQKHFLYCANAPLTFKDGTSIA 539

QY 540 LSCMADYCAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLLAQKLEWAEAFLEE 599
DB 540 LSCMADYCAPVFPFLAIGYKQYTDYSEBALITFSLNNYPAGDPRMAQKLEWAEAFLEE 599

QY 600 MEAFORMMAGMFQVTFATERSLEDEINRTTAEDLPATSYIVIFLYISLAGSYSSWSR 659
DB 600 MESFORNTSDKEQVAFSERSLEDEINRTTIQDLFPVFAVSIIIVFIIYISLAGSYSSCSR 659

QY 660 VMVDSKATLGLGGVAVLGVANMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVL 719
DB 660 VAVESKATLGLGGVIVLGVAVLAANGFYSLGVPSLSVLIQVVPFLVAVGADNIFIFVL 719

QY 720 EYQRLPRPGEPREVHIIRALGRVAPSMLLCSLSAICFFLGAITPMPAVRTFALTSLGA 779
DB 720 EYQRLPRPGEQREAHIGRTLGSVAPSMLLCSLSAICFFLGAITPMPAVRTFALTSLGA 779

QY 780 VILDFLLQMSAFVALLSDSKQEASRLDVCCKVPQELPPQCGEGILLGFFQKAYAPP 839
DB 780 IILDFFLQMTAFVALLSDSKQEASRPDLVCCFSTRKLPPPKKEGELLRRFFRKIYAPP 839

840 LLHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFYLNRIFYEVGAP 899
DB 840 LLHRFIRPVMMLFLTLFGANLYLMCNINVGLDQELALPKDSYLLDYFYLNRIFYEVGAP 899

QY 900 VYFVTTGLGVNFSSEAGMNAICSSAGCNWFSFTQKIQYATEFFEQSYLAIPASSWVDDFID 959
DB 900 VYFVTTSGPNFSSEAGMNAICSSAGCKSFSLTKIQYASEFPDQSYVAIAASSWVDDFID 959

QY 960 WLTP-SSCCRLYISGPNKDKFCPSTVNSLNCNLKNCMSITWGSVRPSVEOFHKYLPWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCNLKNCNRTLGPVRPTAEQFHKYLPWFLND 1019

QY 1019 RNINICPKGGLAAYSTSVNLTSDGVLSRFRWAYHKPLKNSQDYTEALRAARELAANITA 1078
DB 1020 PNINIRCPKGLAAYRTSVNLTSDGVLSRFRWAYHKPLKNSQDYTEALRAARELAANITA 1079

QY 1079 DLRKVPGTDPAPFEPVPTITNVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRSG 1138
DB 1080 DLRKVPGTDPNFEVFPYTIISNVFYQOYLTVLPEGLFTLALCFVPTFVVCYLLGLDMLCSG 1139

QY 1139 LNLLSIVMILVDYTVGFMAWDISYNAVSLINLVSAGMSVFSVSHITRSPFAISTKPTWL 1198
DB 1140 ILNLISILMILVDYTVGLMAVWGISYNAVSLINLVTAVGMSVFSVSHITRSPFAISTKPTWL 1199

QY 1199 ERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITLLGLLHGLVFL 1258
DB 1200 ERAKDATVMSGSAVAGVAMTNFPGLIJLGPAAQLIQIFFRLNLLITLLGLLHGLVFL 1259

QY 1259 PVILSYGDDVNPALALQKRAEBAVAVMWASCPNHSRSTADNIYNHSFEGSI-KG 1317
DB 1260 PVVLSYLGDDVNPQALVQBEKLESEA-AVAPEPSCPYPSPADADAN--VNYGFAPELAHG 1316

QY 1318 AGAISNFLPNNGROF 1332
DB 1317 ANAARSSLPKSDQKF 1331

RESULT 13
US-10-646-301A-2
; Sequence 2, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-646-301A-2

Query Match      78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQLSEPYTTTHQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGEYTPKHEAGVCTFYEECGKPELSGLTSLNVCSLSN 59

QY 61 TPARKITGDHLLILLOKICPRLYTPNTQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
DB 60 TPARHVTGEHALLQRIQCPRLYNGPNTTFACCSKQLLSLESSMSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQAGOLPAWVAEAPYQHSFAEQSYDSCSRVRVPA 179
DB 120 FVSLCHNTCSPPQSLFINVTRVVERGAGEPPAVVAEAFYQKSPAEKAYECSQVRIIPA 179

QY 180 AATLAVGTCMGVYGSALCNAQRLNFOGDTGNGLAPLDITFHLLPEQAVGSIQPLNKG 239
DB 180 AASLAVGSMCGVYGSALCNAQRLNFOGDTGNGLAPLDITFHLLPEQAVGSIQPLNKG 239

QY 240 VARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLGOMPGSLVLIILICLSFVAVT 299
DB 240 IAPCNESQDDSAVSCQDCAASCPVIPPPEALRPSFYMGRRPGWLALIIITFAVFLVS 299

QY 300 ILLVGRVAPARDKGMVDPKXGTSLSDKLSFTHTLLGQFFQGMGTWVASWPLTILVLS 359
DB 300 AVLVRVSVNRNKNKAEGQEPAPKLPHKHKLSPTHTLGRFFQNGTRVVASWPLTILVLS 359

QY 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDHGFPPFRTNQVILTAPNRSS 419
DB 360 FIVWIALAAGLVFTLTDPVELWSAPNSQARSEKSFHDEHGFPPFRTNQVILTAPNRSS 419

QY 420 YRYSLLLPKPNFSGILDLDLLELQERLHLQVWSPAEARNISLQDICYAPLNPYN 479
DB 420 KYYSLLLPKPNFSGILSLDLLELQERLHLQVWSPAEARNISLQDICYAPLNPYN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLGQTSQVDMQKHFLYCANAPLTFKDGTLA 539
DB 480 TSLSDCCVNSLLQYFQNNRTLLMLTANQTLGQTSQVDMQKHFLYCANAPLTFKDGTSIA 539

QY 540 LSCMADYCAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLLAQKLEWAEAFLEE 599
DB 540 LSCMADYCAPVFPFLAIGYKQYTDYSEBALITFSLNNYPAGDPRMAQKLEWAEAFLEE 599

QY 600 MEAFORMMAGMFQVTFATERSLEDEINRTTAEDLPATSYIVIFLYISLAGSYSSWSR 659
DB 600 MESFORNTSDKEQVAFSERSLEDEINRTTIQDLFPVFAVSIIIVFIIYISLAGSYSSCSR 659

QY 660 VMVDSKATLGLGGVAVLGVANMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVL 719
DB 660 VAVESKATLGLGGVIVLGVAVLAANGFYSLGVPSLSVLIQVVPFLVAVGADNIFIFVL 719

QY 720 EYQRLPRPGEPREVHIIRALGRVAPSMLLCSLSAICFFLGAITPMPAVRTFALTSLGA 779
DB 720 EYQRLPRPGEQREAHIGRTLGSVAPSMLLCSLSAICFFLGAITPMPAVRTFALTSLGA 779

QY 780 VILDFLLQMSAFVALLSDSKQEASRLDVCCKVPQELPPQCGEGILLGFFQKAYAPP 839
DB 780 IILDFFLQMTAFVALLSDSKQEASRPDLVCCFSTRKLPPPKKEGELLRRFFRKIYAPP 839
```

```
Qy 120 FVNLHCHNTCSNOSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAESYDSCSRVRPA 179
Db 120 FVSLHCHNTCSPDQSLFINVTRVVERGAGEPPAVVAYEAFYQSFPAEKAYESCQVRIPA 179
Qy 180 AATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVSGIQPLNEG 239
Db 180 AASLAVGSMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQALPDGLOPLNGK 239
Qy 240 VARNESQDDVATCSCODCAASCPAIARPOALDSTFYLGOMPGLVLIILILCSVFAVVT 299
Db 240 IAPCNESQDDSAVCSQDCODCAASCPVIPPPEALRPSFYMGRMPGWLALIIIFTAVFVLLS 299
Qy 300 ILLVGFVAPARDKSKVDPKKTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLS 359
Db 300 AVLVRVVRNKNKAEGQAEAPKLPKHKLSPHTILGRFFQNGWTRVASWPLTILVLS 359
Qy 360 VIPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPARNSS 419
Db 360 FIVVIALAAGLTFTIELTTPDVELWSAPKSQARKEKSFHDEHFGFPFRTNQIVTARNSS 419
Qy 420 YRYDSLILGPKNFGSILDLLELLELQERLRLHQLVWSPQAQRNISLQDICYAPLNPDN 479
Db 420 YKYDSLILGSKNFGSILDLLELLELQERLRLHQLVWSPQAQRNISLQDICYAPLNPYN 479
Qy 480 TSLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDMKDFLYCANAPLTFKDGITALA 539
Db 480 TSLSDCCVNSLLQYFQNNRTLLMTANQTLNGQTSLVDMKDFLYCANAPLTFKDGTSILA 539
Qy 540 LSCMADYGAPVFPFLAIGYKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAEFLKE 599
Db 540 LSCMADYGAPVFPFLAVGQYQTDYSEAEALIIITFLSNYPADDPRLAQAKLWEAEFLKE 599
Qy 600 MRAFORMAGHMQVTFTHAERSLEBENRTTAEIDLPFATSVIIVFLYISLALGYSWSR 659
Db 600 MESFORNTSDKQFAFSAERSLEBENRTTQDLFPVFAVSIIIVFLYISLALGYSYCSR 659
Qy 660 VMVDSKATLGLGGVAVLVGVAAMAAGPFSYLGIRSSLVLOVDFPLVLSVGADNIFIVL 719
Db 660 VAVESKATLGLGGVIVLVGAVLAAMGFYSLGVFSSLVIIQVDFPLVLAUGADNIFIVL 719
Qy 720 EYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALICFFLGLALTPMPAVRTFALTSLGA 779
Db 720 EYQRLPRMPGREGRAHIGRTLGSVAPSMLLCSLSEALICFFLGLALTPMPAVRTFALTSLGA 779
Qy 780 VILDFLLQMSAFVALLSDSKROEASRLDVCCKVPQBELPPPGQGGELLILGFOKAYAPF 839
Db 780 IILDPLQMTAFVALLSDSKROEASRPDLVLCFSTRKLPPLPPKXEGILLRFFRKYAPF 839
Qy 840 LLHWITRGVLLLFALFGVSLYSMCHISVGLDOELAPKDSYLLDYFLFLNRYFEVGP 899
Db 840 LLHRFIRPVWMLFTLFGANLYLNCINVLGDELAPKDSYLLDYFLFLNRYLEVGP 899
Qy 900 VYFVTTLGYNFSSSEAGMNAICSSAGCNNSFTQKIQYATEPPEQSYLEAIPASSWDDPID 959
Db 900 VYFVTTSFNFSSEAGMNAICSSAGCKSFSLTQKIQYASEFPDQSYVAIAASSWDDPID 959
Qy 960 WLTP-SSCCRLYISGPNKDKPCPSTVNSLNCNKCMSITMGSVRPSVEQFHLYLPWFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDEFPCPSTDTSNCLNCKMNRTLGVRPTAEQFHLYLPWFLND 1019
Qy 1019 RPNLTKCPKGLAAVSTSNLTSQGVLSRPMAYHKPLKNSQDYTEALRAARELAANITA 1078
Db 1020 PPNLRCPKGLAAVSTSNLTSQGVLIASQPMAYHKPLRNSQDFTFALRASRLAANITA 1079
Qy 1079 DLKRVPGTDPAFEPYPTITNVFYEQYLTIILPEGLFMLSCLVPTFAVSCILLGLDLRS 1138
Db 1080 DLKRVPGTDPAFEPYPTISNVFYQYLTVLPEGIFTLALCFVPTFVVCYLLGLDMCSG 1139
Qy 1139 LNLLSIYVILVDTVGFMAWDISYNAVSLINLVSACMSVEFVSHITRSFAISTKPTWL 1198
Db 1140 ILNLSIIMILVDTIGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITRSFAVSTKPTRL 1199
```

---

```
Qy 1199 ERAKEATISMGSVAFVAMTNLPGLVLGLAKAQLIQIFFRLLNLLITLLGLHLGLVFL 1258
Db 1200 ERAKDATVFGMSVAFVAMTNFPGLLILGLFAQQLIQIFFRLLNLLITLLGLHLGLVFL 1259
Qy 1259 FVLISYVGPDPVNLPALALEQKRAEEAANAAMVASCNPHSPVSTADNIIYVNHSPGSI-KG 1317
Db 1260 FVLSYLGPDVNQALVQEEKLASEA-AVAPEPSCPQVPSPADADAN--VNYGFAPELAHG 1316
Qy 1318 AGAISNPLPNNGRQF 1332
Db 1317 ANAARSLPKSDQKF 1331

RESULT 14
US-10-736-769-2
; Sequence 2, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Luquan
; APPLICANT: Altmann, Scott W
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-736-769-2

Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAAAGLWLLWALLRLAQSSEYTHQPGYCAFYDECCKNPESLGLTSLSNVSLSN 60
Db 1 MAAAWL-GWLLWALLSAAQGLYTPKHEAGVCTFYEECKNPESLGLTSLSNVSLSN 59
Qy 61 TPARKITGDHLILLQKICPRLYTGPNTO-ACCSAKQIVLSIASISITKALLTRCPACSDN 119
Db 60 TPAPHVTGEHALLQRIICPRLYNGPNTTFACCSTKQLLSLESSISITKALLTRCPACSDN 119
Qy 120 FVNLHCHNTCSNQSFLINVTRVAQLGAGQLPAVVAYEAFYQHSFAESYDSCSRVRPA 179
Db 120 FVSLHCHNTCSPDQSLFINVTRVVERGAGEPPAVVAYEAFYQSFPAEKAYESCQVRIPA 179
Qy 180 AATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVSGIQPLNEG 239
Db 180 AASLAVGSMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQALPDGLOPLNGK 239
Qy 240 VARNESQDDVATCSCODCAASCPAIARPOALDSTFYLGOMPGLVLIILILCSVFAVVT 299
Db 240 IAPCNESQDDSAVCSQDCODCAASCPVIPPPEALRPSFYMGRMPGWLALIIIFTAVFVLLS 299
Qy 300 ILLVGFVAPARDKSKVDPKKTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLS 359
Db 300 AVLVRVVRNKNKAEGQAEAPKLPKHKLSPHTILGRFFQNGWTRVASWPLTILVLS 359
Qy 360 VIPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPARNSS 419
Db 360 FIVVIALAAGLTFTIELTTPDVELWSAPKSQARKEKSFHDEHFGFPFRTNQIVTARNSS 419
```



Db	783	DFLLQMTAFVALLSLDSKRQESRPVVCCFSSENLPFPKQKEGGLLCCFFRKIYTFPLH	842
Qy	843	WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYF	902
Db	843	RFIRPVVLLFLVLFGANLYLMCNISVGLDQDLALPKDSYLLDYFLFLNRYLEVGPVYF	902
Qy	903	VTTILGVNFSSEAGNNAICSSAGCNRNFTOKIOYATEFFEQSYLAIPASSWVDDFDWLT	962
Db	903	DTTSGYNFSTEAGNNAICSSAGCESFSLTKIOYASEFPNQSYVAIAASSWVDDFDWLT	962
Qy	963	P-SSCCRLYISGPNKDFCSTVNSLNCNKMCSITMGSVRPSVEQPHKYLPPWFLNDRN	1021
Db	963	PSSSCCRIYTRGPHKDFCFSTDTSFNCLKNCNMTLGPVRPTTEQPHKYLPPWFLNTPN	1022
Qy	1022	IKCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGLAAYRTSVNLSSDQIIASQFMAYHKPLRNSQDTEALRASRLAANITAE LR	1082
Qy	1082	KVPGTDPAFEVPPYTIITNVFEQVLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLN	1141
Db	1083	KVPGTDNFEVPPYTIISNVFIQQYLTIVLPEGIFTLALCFVPTFVVCYLLGLDLRSGLN	1142
Qy	1142	LLSIWMLVDTVGFMAWDISYNAVSLINLVSAVGMSEVFVSHITRSFAISTKPTWLER	1201
Db	1143	LLSIWMLVDTIGLMAVWGISYNAVSLINLVAVGMSVEFVSHITRSFAVSTKPTRLER	1202
Qy	1202	KEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLLITLLGLHGLVFLPVI	1261
Db	1203	KDATIFMGSAVFAGVAMTNFPGLILGLFAQAQLTIQIFFRNLNLLITLLGLHGLVFLPV	1262
Qy	1262	LSYVGPDPNPALEOKRAEAAVAVMVASCPNHPRSVSTADNIYVNHSEGS- IKGAGA	1320
Db	1263	LSYLGPDVNQALVLEEKLATEA-AMVSEPCQPPFPADANTS DYVNGFNDEPIPEINA	1321
Qy	1321	ISNFPNNGRQF	1332
Db	1322	ASSSLPKSDQKF	1333

Search completed: April 7, 2006, 19:37:11  
Job time : 194 secs

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases: older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**THIS PAGE IS BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:34:15 ; Search time 27 Seconds  
(without alignments)  
1538.804 Million cell

**Title:** US-10-736-769-4

**Perfect score:**

Sequence: 1 MAEAGLRGWLLWALLRLAQ.....GSIKGAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

Database : Published Applications AA New:\*

- 1: /SID55/ptodata/1/pubpaa/US08 NEW PUB pep.:
- 2: /SID55/ptodata/1/pubpaa/US06 NEW PUB pep.:
- 3: /SID55/ptodata/1/pubpaa/US07 NEW PUB pep.:
- 4: /SID55/ptodata/1/pubpaa/PCT NEW PUB pep.:
- 5: /SID55/ptodata/1/pubpaa/US03 NEW PUB pep.:
- 6: /SID55/ptodata/1/pubpaa/US10 NEW PUB pep.:
- 7: /SID55/ptodata/1/pubpaa/US11 NEW PUB pep.:
- 8: /SID55/ptodata/1/pubpaa/US60 NEW PUB pep.:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match Length	DB	ID	Description	
1	6896	99.8	1332	7	US-11-242-459-9	Sequence 9, Appl	
2	2402.5	34.8	1278	6	US-10-995-561-952	Sequence 952, Appl	
3	942	13.6	194	7	US-11-242-459-8	Sequence 8, Appl	
4	195.5	2.8	1087	7	US-11-242-459-17	Sequence 17, Appl	
5	195.5	2.8	1392	7	US-11-242-459-40	Sequence 40, Appl	
6	191	2.8	127	7	US-11-096-568A-6778	Sequence 6778, Ap	
7	173	2.5	85	7	US-11-096-568A-6779	Sequence 6779, Ap	
8	143	2.1	445	7	US-11-242-459-35	Sequence 35, Appl	
9	143	2.1	456	7	US-11-242-459-34	Sequence 34, Appl	
10	141.5	2.0	1048	6	US-10-392-234A-18	Sequence 18, Appl	
11	140.5	2.0	1046	6	US-10-392-234A-16	Sequence 16, Appl	
12	134	1.9	1067	6	US-10-467-657-2648	Sequence 2648, Ap	
13	127	1.8	1048	6	US-10-392-234A-14	Sequence 14, Appl	
14	126.5	1.8	200	7	US-11-242-459-16	Sequence 16, Appl	
15	126	1.8	801	6	US-10-793-626-2020	Sequence 2020, Ap	
16	121.5	1.8	1043	6	US-10-392-234A-34	Sequence 34, Appl	
17	120.5	1.7	2426	7	US-11-203-806A-11	Sequence 11, Appl	
18	119	1.7	1048	6	US-10-392-234A-20	Sequence 20, Appl	
19	117.5	1.7	1049	6	US-10-392-234A-12	Sequence 12, Appl	
20	115.5	1.7	914	7	US-11-072-512-2923	Sequence 2923, Ap	
21	113	1.6	1034	6	US-10-392-234A-30	Sequence 30, Appl	
22	112.5	1.6	697	7	US-11-082-389-362	Sequence 362, App	
23	109.5	1.6	758	7	US-11-096-568A-30412	Sequence 30412, A	
24	109.5	1.6	772	7	US-11-096-568A-30411	Sequence 30411, A	
25	109.5	1.6	827	7	US-11-096-568A-30410	Sequence 30410, A	

## ALIGNMENTS

## RESULT 1

Query Match	99.8%	Score 6896;	DB 7;	Length 1332;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1330: Conservative	0;	Mismatches	2;	Indels 0;
Gaps	0;			

Qy	1	MAEAGLRGWLIIWALLRLAQS	EPYTTTHQPGYCAFYDEC	GNPBLSGSLMTLSNV	CSLUN 60
Db	1	MAEAGLRGWLIIWALLRLAQS	EPYTTTHQPGYCAFYDEC	GNPBLSGSLMTLSNV	CSLUN 60
Qy	61	TPARKITGDHILILLOKICPR	LYTGPNTQACCSAKQLVSL	EASISITKALLTRCPAC	SDNF 120
Db	61	TPARKITGDHILILLOKICPR	LYTGPNTQACCSAKQLVSL	EASISITKALLTRCPAC	SDNF 120
Qy	121	VNLHCHNTCSPNQSLFNV	TRVRAQLGAGOLPAVVAY	EAFYQHSFAEQSYDSC	SRVRVPAA 180
Db	121	VNLHCHNTCSPNQSLFNV	TRVRAQLGAGOLPAVVAY	EAFYQHSFAEQSYDSC	SRVRVPAA 180
Qy	181	ATLAVGTMCGVGSALCNA	ORWLNFGDGTGNGLA	PLDITFHLLPEGQAVSG	GIGPLNEG 240
Db	181	ATLAVGTMCGVGSALCNA	ORWLNFGDGTGNGLA	PLDITFHLLPEGQAVSG	GIGPLNEG 240
Qy	241	ARCNESQGDVATCSQDCA	ASPAIARPAQLDSTFY	LGMQPSGLVIIILCSF	FAVVTTI 300
Db	241	ARCNESQGDVATCSQDCA	ASPAIARPAQLDSTFY	LGMQPSGLVIIILCSF	FAVVTTI 300

Qy	301	LLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGTGWASWPLTILVLSV	360
Db	301	LLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGTGWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTELTDVPELWASPNQARSSEKAFHDQHFGRFTNQVILTPAKRSSY	420
Db	361	IPVVALAAGLVFTELTDVPELWASPNQARSSEKAFHDQHFGRFTNQVILTPAKRSSY	420
Qy	421	RYDSLILGPKNFSGILDLLELLELEQLERLHLQVMSPEAQRNLSIQDICYAPLNPNDT	480
Db	421	RYDSLILGPKNFSGILDLLELLELEQLERLHLQVMSPEAQRNLSIQDICYAPLNPNDT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGOTAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGOTAL	540
Qy	541	SCMADYGAPVPPFLAIGGYKGDYSEABALIMTFSLNYPAGDPRLAQAKLWBEAFLEEM	600
Db	541	SCMADYGAPVPPFLAIGGYKGDYSEABALIMTFSLNYPAGDPRLAQAKLWBEAFLEEM	600
Qy	601	RAFQRMAGMFOVTFTAERSLEDEINRTAEDLPFATSVIVIFLYISLALGSYSMSRV	660
Db	601	RAFQRMAGMFOVTFMAERSLEDEINRTAEDLPFATSVIVIFLYISLALGSYSMSRV	660
Qy	661	MVDSKATLGLGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE	720
Db	661	MVDSKATLGLGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE	720
Qy	721	YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSLGLAV	780
Db	721	YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSLGLAV	780
Qy	781	ILDFFLQWASAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGLLLGFFQKAYAPFL	840
Db	781	ILDFFLQWASAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGLLLGFFQKAYAPFL	840
Qy	841	LHWITRGVWLLFLFALFGVLSYMSCHISVGLDOELALPKDSYLDYDFLFLNRYEVEGAPV	900
Db	841	LHWITRGVWLLFLFALFGVLSYMSCHISVGLDOELALPKDSYLDYDFLFLNRYEVEGAPV	900
Qy	901	YFVTTILGYNFSEAGNNAICSSAGCNFSFTQIOYATEFPFQSYLAIPASSWDDFDW	960
Db	901	YFVTTILGYNFSEAGNNAICSSAGCNFSFTQIOYATEFPFQSYLAIPASSWDDFDW	960
Qy	961	LTPSSCCRLYISGPNKDFCPTVNSLNCNMSITMGSVRPSVEQFHKYLPMFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCPTVNSLNCNMSITMGSVRPSVEQFHKYLPMFLNDRP	1020
Qy	1021	NIKCPKGGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL	1080
Db	1021	NIKCPKGGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAFVFPYITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL	1140
Db	1081	RKVPGTDPAFVFPYITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL	1140
Qy	1141	NLLSIVMILVDVTFGMALWDISYNVAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER	1200
Db	1141	NLLSIVMILVDVTFGMALWDISYNVAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER	1200
Qy	1201	AKEATISMGSAVPAGVAMTNLPCILVLGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV	1260
Db	1201	AKEATISMGSAVPAGVAMTNLPCILVLGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV	1260
Qy	1261	ILSYGPDVNPALALQKRAEAAVAVMVASCPNHPHSRVSTADNIYVNHSPFGS1KGAGA	1320
Db	1261	ILSYGPDVNPALALQKRAEAAVAVMVASCPNHPHSRVSTADNIYVNHSPFGS1KGAGA	1320
Qy	1321	ISNFLPNNGROF	1332
Db	1321	ISNFLPNNGROF	1332

RESULT 2

US-10-995-561-952

; Sequence 952, Application US/10995561

; Publication No. US2005027054A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CL001559

; CURRENT APPLICATION NUMBER: US/10/995,561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 952

; LENGTH: 1278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-995-561-952

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;

Best Local Similarity 39.8%; Pred. No. 2.6e-181;

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy	7	RGWLLWALLRLAQSEPTTTHQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS	59
Db	4	RGLALGLLLLLCPAQVFSQ-----SCVWYEGGIAYGDKRYNCEYSG-----	46
Qy	60	NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDN	119
Db	47	--PPKLPKDGVDLVQELCPGFFFG--NVSLCCDVRLOLTKONLQPLQFLGRCPSCFTN	103
Qy	120	FVNHLCHNTCSNPQSLFINVTR----VAOLGAGQLPAVVAYEAFYQHSFAEOSYDSCSRV	175
Db	104	LLNLFCELTCSRQSOFLNWTATEDYVDPVTNQTKTNVAKELQYVYVQSSPANAMYNACRDV	163
Qy	176	RVPAARATLAVGTWCGYVGSALCNAQRWLNFGQDGTGNGLAPLDIT-----FHLLPQQA	228
Db	164	EAPSSNDKALGLLCKGQADA--CNATNWIEMFNKNGOAPFTTTPVFSDFPVH-----	215
Qy	229	VGSGIQPLNBEVARNCHESQGDVATCSODCAASCAPAIARPO-----ALDSTFYLQ	279
Db	216	---GMEPMNATKGCDESDEVTAPCSQCDSIVCGPKQPPPPPPAPWTILGLDAMYIM	272
Qy	280	QMPGSLVLLIILCSVPAVVTILL-----VGRVAPARDKSMVDPKKGTGTSLS	326
Db	273	WITYMAFLLVFFGFAFVAVWCYKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC	324
Qy	327	DKLSFSTHTLLGQFFQCGTGWASWPLTILVSVIPVVALAAGLVFTELTDVPELWASAP	386
Db	325	DPVSAAPFEGCLRLRFTRWGSCFVRNPGCVIFFSLVFTTACSSGLVFRVYRTNTPVDLWASAP	384
Qy	387	NSQARSEKAFHDQHFGRFTNQVILTPAKRSSYRYSLSLLGPK--NFSGLDLDLLELL	445
Db	385	SSQARLSEKYPDQHFGRFTTEQLITRAPLTDKHIYQPYPSGADVPGPPLDIQLHQVL	444
Qy	446	ELQERLHLQVMSPEAQRNLSIQDICYAPLNPNTSYDCCINSLLQYFQNNRTLLLLTA	505
Db	445	DLQTAIEN--ITASVDNETVTLQDICALPLSPYNT--NCTILSVLNYFQNSHVLDRHK	499
Qy	506	NOTLMGQTSQVDWKDHFPLYCANAPLTFKDGOTALALSCADYAGVFPFLAIGGYKGDYS	565
Db	500	GDDPFF--VYADYHTEFLYCVRAPASLNDTSLLDHPCLTGTFGGVFPFLVGGGDDQNYN	556
Qy	566	EAEALIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMFOVTFTAERSLEDEI	625
Db	557	NATALVITFPNNYNDTEKLQRAQANEKEFINFKYKN---PNLITSFTAERSIEDEL	613
Qy	626	NRTTAEDLPFATSVIVIFLYISLALGSYSMSRWMSVDKATLGLGVAVVLGAVMAAGM	685
Db	614	NRESDSVFTVVISYAINFLYISLALGHKICRRLVDSKLSGLTAGILVLSVACSLG	673

QY 686 FFSYLGIRSSVLVQVPPFLVSVGADNIFIFVLEYQRLPRRPGEPREVIHGRALGRVAP 745  
Db 674 VFSYIGLPLFLIIVIFPLVAVGVDNIFILYQAYQDERLQOETLDOQLGRVLGEVAP 733  
QY 746 SMLCSSLSEACFFFLGALTMPMPAVRTPALTSLGLAVILDFLLQMSAFVALLSLDKRQES 805  
Db 734 SMLFSFSEVAIFLGNLSWMPAVHTFSLFAGLAVIFDLQITCFVSLGLGLDIKKQEK 793  
QY 806 RLDVCCVKPQELPPQ--GEGLLGFFQKAYAPFFLLHWITRGVLLFLFLFVGLSYSM 864  
Db 794 RLDIFCCVRCAGDGTSSQASESCLFRFFKNSYSPLLKQWMPDIVIAIFVGLVSPSLAVL 853  
QY 865 CHLSVGLDQELALPKDSYLLDDYFLFLNRYFEVAGPVYFVTTLGNFSSSEAGMAICSSAG 924  
Db 854 NKVDIGLDQSLSPDDSYMYDYFKSISQYLHAGPPVYFVLEEGHYDTSSKQGNWVCGMG 913  
QY 925 CNNFSTOKIOYATEPEQSYLAI PASSWVDDFIDMLTP--SSCCRLYISGPNKDKCPST 983  
Db 914 CNNDLSVQQLFNAAQLDNRYTRIGFAFSSWIDDYFDWVKPOSSCCRV---DNITDQFCNAS 970  
QY 984 VNSLNCILKNCMSIT--MGSVRPSVEQFHKLPLFWFLNDRPNIKCPKGGLAAYSTSVN--LTS 1040  
Db 971 VVDPACVR--CRPLTPGKQRPQGGDFMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029  
QY 1041 DGOVLASRFWAYHKPLKNSQDYTEALRAARELAANTADLRKVPGTDPAPFVPPYITNV 1100  
Db 1030 GTRVGATYFTWYITVLQTSADFIDALKKARLIASNTV--ETWINGIS--AYRVPPYSVFVY 1086  
QY 1101 FBYOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYVILVDTVGFMAWD 1160  
Db 1087 FBYOYLITLIDDTIFNLGVSLGALFLVTWVLLGCELMSAVIMCATIANVLNMFVGNWMLG 1146  
QY 1161 ISTNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLBRAKEATISMGSAVPAGVAMTN 1220  
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKGSVERAEAEALAHMGSSVFGITLTK 1206  
QY 1221 LPOGILVLGAKAQLIOIFFRMLNLLTLGLLHGLVFLPVILSYVGPVNPA 1272  
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVNKA 1258

RESULT 3  
US-11-242-459-8  
; Sequence 8, Application US/11242459  
; Publication No. US20060035835A1  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: SATOMI, Tomoko Komiyama  
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/11/242,459  
; PRIOR FILING DATE: 2005-10-03  
; PRIOR APPLICATION NUMBER: US/10/239,316  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 8  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Human  
US-11-242-459-8

Query Match 13.6%; Score 942; DB 7; Length 194;  
Best Local Similarity 99.5%; Pred. No. 5.1e-67;  
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 615 FTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSSYSSWSRVMDSKATLGLGGVA 674  
Db 1 FMAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSSYSSWSRVMDSKATLGLGGVA 60

QY 675 VVLGAVMAAGFFSYLGRSSVLVQVPPFLVSVGADNIFIFVLEYQRLPRRPGEPREV 734  
Db 61 VVLGAVMAAGFFSYLGRSSVLVQVPPFLVSVGADNIFIFVLEYQRLPRRPGEPREV 120  
QY 735 HIGRALGRVAPSMMLCSSLSEACFFFLGALTMPMPAVRTPALTSLGLAVILDFLLQMSAFVAL 794  
Db 121 HIGRALGRVAPSMMLCSSLSEACFFFLGALTMPMPAVRTPALTSLGLAVILDFLLQMSAFVAL 180  
QY 795 LSLSKQREASRLD 808  
Db 181 LSLSKQREASRLD 194

RESULT 4  
US-11-242-459-17  
; Sequence 17, Application US/11242459  
; Publication No. US20060035835A1  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: KITA, Shunbun  
; APPLICANT: SATOMI, Tomoko Komiyama  
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/11/242,459  
; PRIOR FILING DATE: 2005-10-03  
; PRIOR APPLICATION NUMBER: US/10/239,316  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 17  
; LENGTH: 1087  
; TYPE: PRT  
; ORGANISM: Human  
US-11-242-459-17

Query Match 2.8%; Score 195.5; DB 7; Length 1087;  
Best Local Similarity 19.5%; Pred. No. 6.1e-07;  
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;  
QY 576 LNNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMFQVTTFAERSLED----- 623  
Db 100 LPNYISVDDR-----WEE-----QRAKFSFVTVYVAMLAQSTSKVQLYGGTD 144  
QY 624 ----EINRTAED--LPIPATSYIVIFLYISALGSSYSSWSRVMDSKATLGLGG--VAVV 676  
Db 145 LPDYEVRRTFNNDMLLAFITSSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALF 199  
QY 677 LGAVMAAGFFSYLGRSSVLVQVPPFLVSVGADNIFIFVLEYQRLPR--RPGEPREVIH 735  
Db 200 LYHVVFGLI---QYLGILNG-----VAAFVIVGIVGDDVFVINTYRQATHLEDPOQRUMIH 251  
QY 736 IGRALGRVAPSMMLCSSLSEACFFFLGALTMPMPAVRTPALTSLGLAVI---LDLFLQMSAFV 792  
Db 252 TVQTAGK---ATFTTSLTTAAAYANVFSQIPAVHDFGLFMSLIYSCCWLAVLVTMPAL 308  
QY 793 ALLSLDSKQREASRLDVC--CCVKPQELPPQ-- 823  
Db 309 GLWSLYLAPLESSCQTSCHQNCSTSLHFPDGFPAPEQVGGSPAQGPPIYLDLDDIPLL 368  
QY 824 -----GEGLLGFFQKAYAP-----FLAHWI-----TRGVVL 850  
Db 369 EVEEPVSLLELGDVSLVSVSPEGLQFASNTSGRHLIYVOLQELHLHHVLMWSAVKSRWVIV 428  
QY 851 LILFLALFGVSL-----YSM-----CHTSVGLDQD- 874  
Db 429 GLFVSLILSLVFPASRLPASAPLFRPDNIQVLLDLKYNLSAEGISCITCSGLFQEK 488  
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFEVAGPVYFVTTLGYNF 910

Db 489 PHSLONNIRTSLEKRRGSGVPWASREATLQDFPGTVVYSKVKQCHPA--VYRLSLNA 546  
Qy 911 SSEAGNVAICSSAG--CNF-----SFTQKIQVATEPEPSYLAIPASSWDDFDIWL 961  
Db 547 SLAPAPQAVSPGDGEVPSQVYRAPFGNFTKLTACMTVGLLQAASPSRKW-----L 600  
Qy 962 TPSSC-----CRLYIS--GP 974  
Db 601 TTLACDAKRGKWFDSFYVATYEQHTRKLYFAQSHKPPFHRGVCMAPPGCLSSSPDGP 660  
Qy 975 NKDKFCPTVNSLNCNKMSITMG-----SVRPSVEQFHKLWPF-----LN- 1017  
Db 661 TKGFF---FVPSKVPKARLSATFGFNPVCVNTGCGKPAVRPLVDTGAMVVFVGIIGNR 717  
Qy 1018 ----DRPNIKCPKGLAAYSTSVNLTS-----GQVL 1045  
Db 718 TRQVDNHVIGDP--GSVVYDSSFDLFKEIGHLCHLCKAIAANSELVKPGQAQCLPSGYSI 775  
Qy 1046 ASRFMAYHKPLKNSOD-----YTEALRAARELAANITADLRKVPDTPAPEVF-P 1094  
Db 776 SSFLQMLHPECKELPEPNLLPGQLSHGAVGVRGVQWISMAFESTTYKGS-SFQTYSD 834  
Qy 1095 YTTINVFYEQYLTIPEG-----LFM-----LSLCL--VP 1122  
Db 835 YLRWESFLOQLOALPEGSVLRRGFTCEHMKQIFMEIVGVQSALCGLVLSLLICVAAVA 894  
Qy 1123 TFAVSCLLGLDLRSLGNLLSIVMLVDVTGFMALWDISYNVNSLINLVSAGVMSVEFV 1182  
Db 895 VFTTHILL-----LLPVLSILGIVCLV--TIMYSGWEM--GAVEAISILVSGSSVDYC 947  
Qy 1183 SHITRSFAIS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI----- 1224  
Db 948 VHLVEGYLLAGENLPPHQAEDARTQW--RTLEAVRHVGVAIVSSALTIVATVPLFFC 1005  
Qy 1225 LVGLAKAQLOIQIFFRNLMLTLGLLHGLVFLPVILSVYGVDPVNPALALEQKRAE--E 1282  
Db 1006 IIAFPK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK 1053  
Qy 1283 AVAAMVVA 1290  
Db 1054 ALGAVLLA 1061

## RESULT 5

US-11-242-459-40  
; Sequence 40, Application US/11242459  
; Publication No. US20060035835A1  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: KITA, Shunbun  
; APPLICANT: SATOMI, Tomoko Komiya  
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/11/242,459  
; CURRENT FILING DATE: 2005-10-03  
; PRIOR APPLICATION NUMBER: US/10/239,316  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 40  
; LENGTH: 1392  
; TYPE: PRT  
; ORGANISM: Human  
US-11-242-459-40

Query Match 2.8%; Score 195.5; DB 7; Length 1392;  
Best Local Similarity 19.5%; Pred. No. 8.6e-07;  
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

Qy 576 LNNYPAGDPLRAQAKLWEEAFLEEMRAFORRMAGHFQVTFTAERSLED----- 623

Db 405 LPNTYISVDDR-----WEE-----ORAKFQSFVTVYVYVLAIAKQSTKSVQVLYGSTD 449  
Qy 624 ----EINRTAED--LPIFATSYIVIFLYXISLALGSSYSSSRVMDVKATLGLGG--VAUV 676  
Db 450 LFDYEVARTFNMDMLAFISSSCIAALVYILTSCSVLSPFGI-----ASIGLSCLVALF 504  
Qy 677 LGAVMAAMGFYSYIGIRSSLVILQVVPFLVLSVGADNIPFVLEYORLPR--RPGEPREHV 735  
Db 505 LYHVVFGI--QYLGIILNG-----VAAFVIVGIVDQDFVFINTRYQATHLEDQLRMWH 556  
Qy 736 IGRALGRVAPSMLLCSLSEALCFPLGALTTPMAVRTALTSLGLAVI---LDFFLOMSAFV 792  
Db 557 TVQTAGK--ATFTSTUTAAVAAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 613  
Qy 793 ALLSLDSKROEASRLDVC--CCVKPQELPPPGQ----- 823  
Db 614 GLWSLYLAPLESSCOTSCONCSRKTSLHFGDVFATPEQVGGSPAQGPPIYLDLDDIPLL 673  
Qy 824 -----GEGLLGFFQKAYAP-----FLLHWI-----TRGVVL 850  
Db 674 EVEEPPVSLGDLGVLSVSPGELQASNTGSRGHLIVQLQELLHHWVLSAVKSRWVIV 733  
Qy 851 LLFLALFGVSL-----YSM-----CHISVGLDOE- 874  
Db 734 GLFVSIILSLVFAURLPASRAPLLFRPDNTIQVLLDLKYNLSAEGISCTICSGLFQEK 793  
Qy 875 -----LALPK-----DSYLLDY--FLFLNRYFEFGAPVYFVTTILGYNF 910  
Db 794 PHSLONNIRTSLEKRRGSGVPWASREATLQDFPGTVVYSKVKQCHPA--VYRLSLNA 851  
Qy 911 SSEAGNVAICSSAG--CNF-----SFTQKIQVATEPEPSYLAIPASSWDDFDIWL 961  
Db 852 SLAPAPQAVSPGDGEVPSQVYRAPFGNFTKLTACMTVGLLQAASPSRKW-----L 905  
Qy 962 TPSSC-----CRLYIS--GP 974  
Db 906 TTLACDAKRGKWFDSFYVATYEQHTRKLYFAQSHKPPFHRGVCMAPPGCLSSSPDGP 965  
Qy 975 NKDKFCPTVNSLNCNKMSITMG-----SVRPSVEQFHKLWPF-----LN- 1017  
Db 966 TKGFF---FVPSKVPKARLSATFGFNPVCVNTGCGKPAVRPLVDTGAMVVFVGIIGNR 1022  
Qy 1018 ----DRPNIKCPKGLAAYSTSVNLTS-----GQVL 1045  
Db 1023 TRQVDNHVIGDP--GSVVYDSSFDLFKEIGHLCHLCKAIAANSELVKPGQAQCLPSGYSI 1080  
Qy 1046 ASRFMAYHKPLKNSOD-----YTEALRAARELAANITADLRKVPDTPAPEVF-P 1094  
Db 1081 SSFLQMLHPECKELPEPNLLPGQLSHGAVGVRGVQWISMAFESTTYKGS-SFQTYSD 1139  
Qy 1095 YTTINVFYEQYLTIPEG-----LFM-----LSLCL--VP 1122  
Db 1140 YLRWESFLOQLOALPEGSVLRRGFTCEHMKQIFMEIVGVQSALCGLVLSLLICVAAVA 1199  
Qy 1123 TFAVSCLLGLDLRSLGNLLSIVMLVDVTGFMALWDISYNVNSLINLVSAGVMSVEFV 1182  
Db 1200 VFTTHILL-----LLPVLSILGIVCLV--TIMYSGWEM--GAVEAISILVSGSSVDYC 1252  
Qy 1183 SHITRSFAIS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI----- 1224  
Db 1253 VHLVEGYLLAGENLPPHQAEDARTQW--RTLEAVRHVGVAIVSSALTIVATVPLFFC 1310  
Qy 1225 LVGLAKAQLOIQIFFRNLMLTLGLLHGLVFLPVILSVYGVDPVNPALALEQKRAE--E 1282  
Db 1311 IIAFPK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK 1358  
Qy 1283 AVAAMVVA 1290  
Db 1359 ALGAVLLA 1366

RESULT 6

```
US-11-096-568A-6778
; Sequence 6778, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6778
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(127)
; OTHER INFORMATION: Ceres Seq. ID no. 15168324
US-11-096-568A-6778

Query Match      2.8%; Score 191; DB 7; Length 127;
Best Local Similarity 44.3%; Pred. No. 6.4e-08;
Matches 35; Conservative 21; Mismatches 23; Indels 0; Gaps 0;

QY 1199 ERAKEATISGSAVFAGVAMTNLPGLVLGLAKAQLIQIFPRLNLLITLLGLHLGLVFL 1258
DB 34 QRAKTALCTMGASVFSGITLTKLVGLVLCFSTSQIFVYVQWYLMALVLGLHLGLVFL 93

QY 1259 PVILSYGPDVNPALALEQ 1277
DB 94 FVVLSLFGPPLRYTVIKEQ 112

RESULT 7
US-11-096-568A-6779
; Sequence 6779, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6779
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(85)
; OTHER INFORMATION: Ceres Seq. ID no. 15168325
US-11-096-568A-6779

Query Match      2.5%; Score 173; DB 7; Length 85;
Best Local Similarity 44.3%; Pred. No. 9.6e-07;
Matches 31; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 1208 MGSVAFGAVAMTNLPGLVLGLAKAQLIQIFPRLNLLITLLGLHLGLVFLPVILSYGVP 1267
DB 1 MGSVAFSGITLTKLVGLVLCFSTSQIFVYVQWYLMALVLGLHLGLVFLPVILSLFGP 60

QY 1268 DYNPALALEQ 1277
DB 61 PLRYTVIKEQ 70

RESULT 8
US-11-242-459-35
; Sequence 35, Application US/11242459
; Publication No. US20060035835A1
```

```
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 35
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-35

Query Match      2.1%; Score 143; DB 7; Length 445;
Best Local Similarity 25.4%; Pred. No. 0.0024;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

QY 576 LNNYPAGDPRLAQAQLWEAFLEEMRAFQRRMAGMFQVTFTAERSLED----- 623
DB 100 LPNYSVDDR-----WEE-----QRAKEQSFVVTYVAMLAQSKVQVLYGGTD 144

QY 624 ----EINRTABD--LPIFATSYVIVFLYISALGSYSSKSRVMDSKATLGLGG-VAVV 676
DB 145 LFDYEVRRTFNNDMLIAFISSCIAALVILTSCSVLFSFFGI-----ASIGLSCLVALF 199

QY 677 LGAVMAAMGFPSYLGIRSSLVILQVVPFLVLSVAGDNIFVLEYQRLPR-RGEPREVH 735
DB 200 LYHVVEGI---QYLGLING-----VAAFVIVGVDVDFVINTYRQATHLEDPPQLRMH 251

QY 736 IGRALGRVAPSMLLCSLSEAICFFLGALTTPMPAVRTFALTSGIAVI---LDFFLLQMSAFV 792
DB 252 TVQTAGK---ATFFSTLTAAAYANVFSQIPAVHDFGLPMSLIVSCCWLAVLVTMPAAL 308

QY 793 ALLSLDSKQERASRLDVC---CCVKPQELPPPG 822
DB 309 GLWSLYLAPLESSCOTSCHQNGSRKTSLHFPG 340

RESULT 9
US-11-242-459-34
; Sequence 34, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 34
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-34

Query Match      2.1%; Score 143; DB 7; Length 456;
Best Local Similarity 25.4%; Pred. No. 0.0025;
```



Db 329 TTPVVSASIH-EVVKTLGE-----AIIIVFLVMYLFQNF-----RATLIPTI 370  
Qy 671 GGVAIVLG--AVMAAGFFSYLGIRSLVILQVVPFLVSVG--ADNIFIVLEYORLPR 726  
Db 371 AVPVLLGTGVLAAFGF-----SINTLTWFG-WVLAIGLLVDDAIVVVENVERVMA 421  
Qy 727 RPG-EPREV---HIGRALGR-VAPSMLLCSLSAICFFIGALTTPMAVRTFALTSGLAVI 781  
Db 422 EELSPREARKSMGOIQGALVGIAMVLSAVFLPMAPFGGSTGVI--YRQFSITIVSAMA 479  
Qy 782 LDELLQWSAFVALLSLDSKQEAERLDVCCC--VKPOELPPPGQEGLLGPFQKAYAP 838  
Db 480 L-----SVTVALI-----LTPALCATMLAPIEKGDHGEHKGFFGFWNFRNPLS 522  
Qy 839 FLHWTIRGVV-----LLFLALFGVSLYSMCHISVGL-----DQ-----ELALPKD 880  
Db 523 -TTHGYERGVASILKRAPYLLIYVIVVAGMIWMTETRIPTAPLPDEDQGVLFQAQVTPPG 581  
Qy 881 S-----YLLDYFLNRYFEVGAIVYFVTTLGYNFSSE----- 913  
Db 582 SSAERTQVVVDSMREYLLEK-----ESSSVSVFTVTGTFNAGRGQSSGMAPTMLXP 633  
Qy 914 -----AGMAI---CSSAGCNNEFTQKIQYATEPEQSYLAIPASSWDDDFD----- 959  
Db 634 WEERPGSENSVELAKRAQHFFSFKDAWVFA--FAPPSVLELGNATGDFLQDOAGVG 691  
Qy 960 ----WLTSSCCRLYISGPNKDKFCPSTVN-----SLNCLKNCMSI 996  
Db 692 HEVLLQARNKFLMLAAQNPAQRVRPNGMSDEPOYKLEIDDEKASALGVSLADINSTVSI 751  
Qy 997 TWGSV-----RPSVEQFHKLWFLNDRPN-----IKCPKGGLA---AYST 1034  
Db 752 AWGSSVYNDPDRGRVKRV-----YLOGRDARMNPDDL SKWYVRNDRKGEMVPPNAPAT 805  
Qy 1035 -----SVNLTSQGVLASRPMAYHKPLKNSODYTEALRAARELANITADLRKVPCTDP 1088  
Db 806 GKWEYGPXLERVNGVPAMEILGEPAPGLSSGD---AMAAVEE---IVKQLPKGVG--- 855  
Qy 1089 AFEVFYTTITVPFYEQLT-----ILPEGLFMLSIC-----LVPTFAVS 1127  
Db 856 -----YSWTGLSYEERLSSGQAPALYALSLLVVFLCLAALYESWSIPFSVMLVPLGVI 909  
Qy 1128 CULLGLDLSGLNLLSIVMILVDTVGFMALWDISNVAVSLNL-----V 1172  
Db 910 GALLATSMR-GLSNDVFFVGLLTITGLSA-----KNAILIVEFAKELHEQKGIVEAAI 963  
Qy 1173 SAVGMSVEFVSHITRSFALSTKPTWLERAKEATISGSAVFAGVAMTNLPGILVL-GLAK 1231  
Db 964 EACRMRLRPVMTSLAFILGVVP-----LAISTG-----AGSGSQHAIGTGVIGGMVT 1011  
Qy 1232 AOLIQIIFPRL 1242  
Db 1012 ATVLAIFWVPL 1022

RESULT 12

US-10-467-657-2648  
; Sequence 2648, Application US/10467657  
; Publication No. US20050260581A1

GENERAL INFORMATION:

; APPLICANT: CHIRON SPA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta

TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: Seqwin99, version 1.04

; SEQ ID NO 2648  
; LENGTH: 1067  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-2648

Query Match 1.9%; Score 134; DB 6; Length 1067;

Best Local Similarity 19.4%; Pred. No. 0.043;

Matches 185; Conservative 139; Mismatches 354; Indels 274; Gaps 44;

Qy 452 RHLOW-SPEAQR--NISLQDICYAPLNPDPNTSLDYCCINSLLQYQNNRTLLLLLTANQT 508  
Db 180 RAMRIVWDPKLQNYNLSFADVGS-LSAQNIIQISAGISGLPAV--RGQT---VTVATVT 233  
Qy 509 LMGQTSQVDMKHFLYCANAPLTFKQGTALALSCMADYGAPVFPFLAIGYKGDYSEAE 568  
Db 234 AQQLGTAEEFGNVILRANT-----DGSNIYLKDAKVGLGM-----EDYSST 277  
Qy 569 ALI-----MTFLNNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFQVTFABRSLE 622  
Db 278 RLNGVNTTCMAVMLSN--SGN-AMATAKAVKERLAVLEKYFPQGMS--WKTPTYDTSKFVE 332  
Qy 623 DEINRTAEDLIPATSIYIVIELYISLALGSSYSSRSRVMVDSKATLGLGVAIVLGAAMA 682  
Db 333 ISIEKVIHTLIBAMVLVVFVVMYLFQNI--RYTLPTIIVVPISL---LGGFA----- 379  
Qy 683 AMGFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYORLPRRPGEPREHVHIGRALGR 742  
Db 380 ---FISYMG-SINVLTMFAMILVIGIVVDDAIVVENVERIMAGEGLPPEKATKKAMGQ 435  
Qy 743 VAPSM-----LCSISEAICFFLGALTTPMAVRTFALTSGLAVILDLLQMSAFVALLSL 797  
Db 436 ISGAVIGITAVLISVFPVPLMFSGAAGNI--YKQFALTWASSI-----AFSAFLAL--- 484  
Qy 798 DSKREASRLDVCCCKVQOELPPPGQEGE--LLGFFQKAYAPF-----LHMIT 845  
Db 485 -----TLTPALCATMLKTIPIKGHHEKKGFFGFWNKKFDSWTHGYEGRVAKVLRTF 536  
Qy 846 RGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAIVYFVTT 905  
Db 537 RMVVYIGLAVGV-----FLPMR-----LPTSFLT 563  
Qy 906 LGYNFSSEAGNMAICSSAGC-----NNFSFTQKIQYATEPPE-QSYLAIPASSWDD--- 956  
Db 564 EDQGFV---MVSQVLPAGATKERTDATLAQVTLAKSIPETENIITVSGFSFGSQNM 619  
Qy 957 -----FTDW--LTSSCCRLYIS-----GPNKDKFCPSTVN-----S 986  
Db 620 AMGFALDKDWERTAGSDAVAVAGKLTGMMMTLKDQFGISVVPVPPILGLNGSGLSIN 679  
Qy 987 LNCLKNCMSITWGSVVRPSVEQFHKLWPF-----LNDPRNICKCPKGLAAVYSTVN 1037  
Db 680 LQDRNNTGHTALLAKRNELIQKRASGLFDPSTVRAGGLEDSFQKIDINRAAAAQGIS 739  
Qy 1038 LTSQGVLASRF-MAYHRKPLKNSODYTEALRAARELANITADLRK-----VP-- 1084  
Db 740 PADIRTALASALSSSYSDFPNQGLQRVMVQADEARMQPADILNLTVPNKSVAVPLS 799  
Qy 1085 -----GTD-----PAFEVFPYTTITVPFYEQL----- 1106  
Db 800 TIATVSWENGTEQSVRFNGYPSMKLSASPATGVSTGQAMAAVQKMVDLGGYSGFWGQ 859  
Qy 1107 -----TILPEGLFMLSICLV-----PTFAVSCLLAGL-----DLR 1136  
Db 860 SREAKGGSQTLILYGLAAVAVFLVLAALYESWSIPLAVILVPLGLIGAAAGVTGRNLF 919  
Qy 1137 SGLL-NLLSIVMILVDTVGFMALWDIS--YNVAVSLNLV-----SAVGSVE-----F 1181  
Db 920 EGLGSPSPANDIYFQGVFTVMGLSAKNAIILIEPAKDLQAQGSKSAVEAAEARLRF 979  
Qy 1182 VSHITRSPA--ISTKPTWLERAKEATI--SHGSAVFAGVAMTNLPGILVLGL 1229  
Db 980 RPIINTSPAFILGVVPLVIAAGASSASQRAIGTTTFWGMVLVGTLLSVFLVPL 1031

## RESULT 13

US-10-392-234A-14  
; Sequence 14, Application US/10392234A  
; Publication No. US20050255538A1  
; GENERAL INFORMATION:  
; APPLICANT: Pharmacia and Upjohn Corporation  
; APPLICANT: Buxser, Steven  
; APPLICANT: Poole, Keith  
; APPLICANT: Decker, Douglas  
; APPLICANT: Xiazhi Li  
; TITLE OF INVENTION: Method for Screening for acrab Transporter Family Inhibitors  
; FILE REFERENCE: 6206  
; CURRENT APPLICATION NUMBER: US/10/392,234A  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: US 60/364,935  
; PRIOR FILING DATE: 2002-03-15  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 1048  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-10-392-234A-14

Query Match 1.8%; Score 127; DB 6; Length 1048;  
Best Local Similarity 18.4%; Pred. No. 0.15;  
Matches 147; Mismatches 107; Indels 304; Gaps 34;  
QY 624 EIRNTAEDLPATSYIVIFLYISALGYSWSRVMVDSKATL--GLGGVAVVLG--A 679  
Db 339 EVVKTLVE-----AIIIVFLVMVFLQNF-----RATLIPIAVPVLLGTFA 381  
QY 680 VMAAMGFYGLRSLVLQVVPFLVSVG--ADNIFVLEYORLPRRGPPEVHIG 737  
Db 382 VLAAGF-----SINTLWFG-MVLAIGLLVDDAIVVVENVERVMAEGLPPKEATR 432  
QY 738 RALGR-----VAPSMILCSLSAICFFLGALPMPAPVTFALTSGTGLAVILDFLLQMSAFV 792  
Db 433 KSMGQIQGALVGIAMVLSAVFIPMAFFGSGTGA--YRQFSITVSNALSVLVALIITP 490  
QY 793 ALLSLDSKQERASLDVCCVQBELPPPGQEGLLGFFQKAYAPFLHMI----- 844  
Db 491 ALCA-----TMLKPIQKSHGATTG--FFGWFNRMPDKSTHVTDSVGNILR 535  
QY 845 TRGVLLFLALF-----GVSLYSMCHI SVGLDQ----- 874  
Db 536 STGRYLVLIIVGMAMLFVRLPSFLPDEQGVFL--SMAQLPAGATQERTQKVLDEMT 594  
QY 875 ---LALPKDS---YLLDYFLFLNR----- 892  
Db 595 NYLTKEDNVSEFAVNGFGFAGRGQNTGIAFVSLKDSQRPGEENKVEAITARAMGYF 654  
QY 893 ---YFVGAPVFTVLGYNFSEAGMNAICSSAGCNPFSTQ----- 932  
Db 655 SOIKDAMVFAFLPAIVELGTATGDFE-----LIDQAGLGHEKLTQARNQLFGMVAQ 707  
QY 933 ---KIQVATEPPEQSVLAIPAS-----SWVDDFDMLTP 963  
Db 708 HPDVLTVGRPNGLETPQFKIDIDQEAQALGVISIDINTLGAAGGSGVYNDFID---R 764  
QY 964 SSCCLXYTSGPNKQKFCSTVNSLNCNKMSTITMGSVRPSVQFHYKLPWFLNDRPNIK 1023  
Db 765 GRVKVYIMSEAKYRMLPEDIGK-----WVRGSDQM 797  
QY 1024 CPKGLAAVST-----SVNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELANIT 1077  
Db 798 VP---FSAFSTSRWEYGSPLRERNGLPSLILGOAPGKST---GEAMALMEELAGKLP 851  
QY 1078 ADLRKVPCTDPAFEVFPYTTITNVFEQVLT-----ILPEGLFMLSICL----- 1120  
Db 852 SGI-----GYDWTGMSYQERLSGNQAPALYALISLIVVFLCLAALYESWSIFP 898

QY 1121 -----VPTFAVSCLLGLDLRSGLLNLSIWMILVDTVGFMAWLDISNVAVSLINL----- 1171  
Db 899 SVMLVVPLGVGVGALLAA--TFRGLTNDVYFQVGLLTIGLSA-----KNALLIVEFAKDL 951  
QY 1172 -----VSAGMSVEFVSHITRSPFAISTKPTWLERAKKATISMGSAVPAGVAMT 1219  
Db 952 MEKEGKGLIEATLEAVRMRLRPILMTSLAFTLGMP-----LVISSG-----AGSGAQ 999  
QY 1220 NLPGLVL-GLAKAQLQIPIF 1239  
Db 1000 NAVGTGVMGVMVTATILAIFF 1020

## RESULT 14

US-11-242-459-16  
; Sequence 16, Application US/11242459  
; Publication No. US20060035835A1  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: KITA, Shunbun  
; APPLICANT: SATOMI, Tomoko Komiyama  
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/11/242,459  
; CURRENT FILING DATE: 2005-10-03  
; PRIOR APPLICATION NUMBER: US/10/239,316  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 16  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Human  
US-11-242-459-16

Query Match 1.8%; Score 126.5; DB 7; Length 200;  
Best Local Similarity 25.7%; Pred. No. 0.015;  
Matches 53; Conservative 38; Mismatches 92; Indels 23; Gaps 8;  
QY 612 QVTFTRASLEDEIRNRTAED--LPIFATSYIVIFLYISALGYSWSRVMVDSKATLG 669  
Db 3 QVLYGGTDLFDYEVRTFNNDMLLAFISSCTAALVILTSCSVFLSFFGI-----ASIG 57  
QY 670 LGG-VAVVLGAVNAAMGFFSYLGRSLVLQVVPFLVLSVGADNIFVLEYQRLPR-R 727  
Db 58 LSCLVALLFLVHVVEGI---QYLGLNG---VAAFVIVGIGVDDVVFVINTYRQATHLE 109  
QY 728 PGEPREHVHIGRALGRVAPSMILCSLSAICFFLGALTPMPAPVTFALTSGTGLAVI---LDF 784  
Db 110 DPQLRMHTVQTAGK---ATPFTSLTAAAYANVFSQIPAVHDFGLFMSLIVSCCWLAV 166  
QY 785 LLOMSAFVALLSLDSKQERASLDVC 810  
Db 167 LVTPPAALGLWSLYLAPLESSCQTS 192

RESULT 15  
US-10-793-626-2020  
; Sequence 20, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P034800S  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472



**THIS PAGE IS BLANK**

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case .***

**THIS PAGE IS BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:22:20 ; Search time 569 Seconds  
(without alignments)  
3235.086 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGATSNFLPNNGRQF 1332

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main.\*

- 1: /cgn2\_6/ptodata/1/paa/US060 COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US061 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US073 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US074 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US075 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US076 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US077 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US078 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US079 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US080 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US097 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US099 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US103 COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US104 COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US105 COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US106 COMB.pcp.\*
- 37: /cgn2\_6/ptodata/1/paa/US107 COMB.pcp.\*
- 38: /cgn2\_6/ptodata/1/paa/US108 COMB.pcp.\*
- 39: /cgn2\_6/ptodata/1/paa/US109 COMB.pcp.\*
- 40: /cgn2\_6/ptodata/1/paa/US110 COMB.pcp.\*
- 41: /cgn2\_6/ptodata/1/paa/US111 COMB.pcp.\*
- 42: /cgn2\_6/ptodata/1/paa/US112 COMB.pcp.\*
- 43: /cgn2\_6/ptodata/1/paa/US114 COMB.pcp.\*

- 44: /cgn2\_6/ptodata/1/paa/US600 COMB.pcp.\*
- 45: /cgn2\_6/ptodata/1/paa/US601 COMB.pcp.\*
- 46: /cgn2\_6/ptodata/1/paa/US602 COMB.pcp.\*
- 47: /cgn2\_6/ptodata/1/paa/US603 COMB.pcp.\*
- 48: /cgn2\_6/ptodata/1/paa/US604 COMB.pcp.\*
- 49: /cgn2\_6/ptodata/1/paa/US605 COMB.pcp.\*
- 50: /cgn2\_6/ptodata/1/paa/US606 COMB.pcp.\*
- 51: /cgn2\_6/ptodata/1/paa/US607 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	1	PCT-US03-22467-4
2	6909	100.0	1332	1	PCT-US03-40113-4
3	6909	100.0	1332	1	PCT-US05-01469-4
4	6909	100.0	1332	36	US-10-621-758A-4
5	6909	100.0	1332	36	US-10-646-301A-4
6	6909	100.0	1332	36	US-10-663-208A-4
7	6909	100.0	1332	37	US-10-736-769-4
8	6909	100.0	1332	37	US-10-750-386-4
9	6896	99.8	1332	1	PCT-US01-04098A-1831
10	6896	99.8	1332	32	US-10-239-316-9
11	6896	99.8	1332	32	US-10-258-899A-1831
12	6896	99.8	1332	32	US-10-293-244-1831
13	6896	99.8	1332	42	US-11-218-141-1831
14	6896	99.8	1332	42	US-11-242-459-9
15	6872.5	99.5	1359	1	PCT-US01-04098A-1830
16	6872.5	99.5	1359	1	PCT-US03-22467-4
17	6872.5	99.5	1359	1	PCT-US03-40113-4
18	6872.5	99.5	1359	1	PCT-US05-01469-4
19	6872.5	99.5	1359	1	PCT-US05-27579-22
20	6872.5	99.5	1359	32	US-10-258-899A-1830
21	6872.5	99.5	1359	32	US-10-293-244-1830
22	6872.5	99.5	1359	36	US-10-621-758A-4
23	6872.5	99.5	1359	36	US-10-646-301A-4
24	6872.5	99.5	1359	36	US-10-663-208A-4
25	6872.5	99.5	1359	37	US-10-736-769-4
26	6872.5	99.5	1359	37	US-10-750-386-4
27	6872.5	99.5	1359	42	US-11-218-141-1830
28	6868.5	99.4	1359	31	US-10-170-205E-21686
29	6868.5	99.4	1359	48	US-60-453-050-8956
30	6868.5	99.4	1359	48	US-60-466-412-8956
31	6868.5	99.4	1359	48	US-60-466-412-8956
32	6536	94.6	1344	1	PCT-US01-08631-53052
33	6536	94.6	1344	34	US-10-450-763-53052
34	5421.5	78.5	1331	1	PCT-US03-22467-2
35	5421.5	78.5	1331	1	PCT-US03-40113-2
36	5421.5	78.5	1331	1	PCT-US05-01469-2
37	5421.5	78.5	1331	36	US-10-621-758A-2
38	5421.5	78.5	1331	36	US-10-646-301A-2
39	5421.5	78.5	1331	36	US-10-663-208A-2
40	5421.5	78.5	1331	37	US-10-736-769-2
41	5421.5	78.5	1331	37	US-10-750-386-2
42	5407	78.3	1333	1	PCT-US03-22467-12
43	5407	78.3	1333	1	PCT-US03-40113-12
44	5407	78.3	1333	1	PCT-US05-01469-12
45	5407	78.3	1333	1	PCT-US05-27579-3

ALIGNMENTS

RESULT 1  
PCT-US03-22467-4  
; Sequence 4, Application PC/TUS0322467  
; GENERAL INFORMATION:  
; APPLICANT: Schering Corporation

```
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22467-4

Query Match      100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSILFINVTRVAOLGAGOLPAVVAEYAFYOHSPAEQSYDSCSRVRPAA 180
DB 121 VNLHCHNTCSNQSILFINVTRVAOLGAGOLPAVVAEYAFYOHSPAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCYVGSALCNAQRLNFGQDTGNGLAPLDITFHLBPQAVGSGIQLNEGV 240
DB 181 ATLAVGTMCYVGSALCNAQRLNFGQDTGNGLAPLDITFHLBPQAVGSGIQLNEGV 240
QY 241 ARCHESQDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILICSPFAVVTI 300
DB 241 ARCHESQDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILICSPFAVVTI 300
QY 301 LLVGRVAPADKSKWDPKKGTSLSDKLSFSTHTLLQFFQGGTGWASWPLTILVLSV 360
DB 301 LLVGRVAPADKSKWDPKKGTSLSDKLSFSTHTLLQFFQGGTGWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHDPFPRTNQVILTAPNRSSY 420
DB 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHDPFPRTNQVILTAPNRSSY 420
QY 421 RYDSLLGPKNPGSGILDLDLLELLELRLHQLVNSPEAQNSISLODICVAPLNPDNT 480
DB 421 RYDSLLGPKNPGSGILDLDLLELLELRLHQLVNSPEAQNSISLODICVAPLNPDNT 480
QY 481 SLVDCCLNSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540
DB 481 SLVDCCLNSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGRDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGYKGRDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLAIGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLAIGSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVVLAGVMAAGFFSVLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATLGLGVAVVVLAGVMAAGFFSVLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTFALTSGLAV 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKVQKPELPPOGGEGLLIGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKVQKPELPPOGGEGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900

; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22467-4

Query Match      100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNNFSTQKIQYATEPPEQSYLAIPASSWVDDPIDW 960
DB 901 YFVTTILGYNFSSEAGNNAICSSAGCNNFSTQKIQYATEPPEQSYLAIPASSWVDDPIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFHKYLPWFNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFHKYLPWFNDRP 1020
QY 1021 NIKCPKGLAAVSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAVSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPFAPEFVPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPFAPEFVPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGFMAWMDISYNVSLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIVMLVDTVGFMAWMDISYNVSLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKRATISMGSAVPAVAMTNLPGILVLGLAKAOLIOIFFFFPRMLNLTLLGLLHGLVPLPV 1260
DB 1201 AKRATISMGSAVPAVAMTNLPGILVLGLAKAOLIOIFFFFPRMLNLTLLGLLHGLVPLPV 1260
QY 1261 ILSVVGPDVNPALAEQKRAEEAAVAVMVASCPNHPRSVSTADNIYVNHSPFEGSIKAGA 1320
DB 1261 ILSVVGPDVNPALAEQKRAEEAAVAVMVASCPNHPRSVSTADNIYVNHSPFEGSIKAGA 1320
QY 1321 ISNPLPNNGRQF 1332
DB 1321 ISNPLPNNGRQF 1332

RESULT 2
PCT-US03-40113-4
; Sequence 4, Application PC/TUS0340113
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K3-WI
; CURRENT APPLICATION NUMBER: PCT/US03/40113
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40113-4

Query Match      100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
```



Db 541 SCWADYCAPVPFFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAOKLWEEAFLEEM 600  
Qy 601 RAFORMMAGMFQVTFFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSSRV 660  
Db 601 RAFORMMAGMFQVTFFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSSRV 660  
Qy 661 MVDKATLGLGGVAVLGVAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVLGVAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720  
Qy 721 YQRLPRRPGEPREVIHIGALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAV 780  
Db 721 YQRLPRRPGEPREVIHIGALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAV 780  
Qy 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQBELPPPGQEGILLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQBELPPPGQEGILLGFQKAYAPFL 840  
Qy 841 LHWITRGVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900  
Db 841 LHWITRGVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900  
Qy 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFDW 960  
Db 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFDW 960  
Qy 961 LTPSSCCRLYISGPNKDFCPTVNSLNCNCSITMGSVRPSVEQPHKYLPMFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDFCPTVNSLNCNCSITMGSVRPSVEQPHKYLPMFLNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHPLKNSODYTEALRAARELAANIATDL 1080  
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHPLKNSODYTEALRAARELAANIATDL 1080  
Qy 1081 RKVPGTDPAFVFPYITINVFYEQYITILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140  
Db 1081 RKVPGTDPAFVFPYITINVFYEQYITILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140  
Qy 1141 NLLSVIMLVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200  
Db 1141 NLLSVIMLVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200  
Qy 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQIQTIEFFRLNLLITLLGLLHGLVFLPV 1260  
Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQIQTIEFFRLNLLITLLGLLHGLVFLPV 1260  
Qy 1261 ILSVYGPDPNPALALEQKRAEEAANAAMVAVASCPNHPRSVSTADNIYVNHSPGSIKGAGA 1320  
Db 1261 ILSVYGPDPNPALALEQKRAEEAANAAMVAVASCPNHPRSVSTADNIYVNHSPGSIKGAGA 1320  
Qy 1321 ISNFLPNNGRQF 1332  
Db 1321 ISNFLPNNGRQF 1332

## RESULT 4

US-10-621-758A-4  
; Sequence 4, Application US/10621758A  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JB01603-K-US  
; CURRENT APPLICATION NUMBER: US/10/621,758A  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 1332

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-621-758A-4  
Query Match 100.0%; Score 6909; DB 36; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MABAGLUGWLLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60  
Db 1 MABAGLUGWLLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60  
Qy 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120  
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180  
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180  
Qy 181 ATLAGVTMCQVYGSALCNAORWLNFGDTCNGLAPLDITFHLLPQGAQVSGIQLNEGV 240  
Db 181 ATLAGVTMCQVYGSALCNAORWLNFGDTCNGLAPLDITFHLLPQGAQVSGIQLNEGV 240  
Qy 241 ARCNESQGDVATCSCQDCCAASCPAIAAPQALDSTFYLGOMPGSLVLIILCSVFAVVTI 300  
Db 241 ARCNESQGDVATCSCQDCCAASCPAIAAPQALDSTFYLGOMPGSLVLIILCSVFAVVTI 300  
Qy 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQPFQGWGTWVASWPLTILVLSV 360  
Db 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQPFQGWGTWVASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSKAHFHQHGFPPFTNOVLLTAPNRSY 420  
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSKAHFHQHGFPPFTNOVLLTAPNRSY 420  
Qy 421 RYDSSLIGKPNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSLQDICYAPLNDPT 480  
Db 421 RYDSSLIGKPNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSLQDICYAPLNDPT 480  
Qy 481 SLVDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKOHFLYCANAPLTFKDGTAIAL 540  
Db 481 SLVDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKOHFLYCANAPLTFKDGTAIAL 540  
Qy 541 SCWADYCAPVPFFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAOKLWEEAFLEEM 600  
Db 541 SCWADYCAPVPFFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAOKLWEEAFLEEM 600  
Qy 601 RAFORMMAGMFQVTFFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSSRV 660  
Db 601 RAFORMMAGMFQVTFFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSSRV 660  
Qy 661 MVDKATLGLGGVAVLGVAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVLGVAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720  
Qy 721 YQRLPRRPGEPREVIHIGALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAV 780  
Db 721 YQRLPRRPGEPREVIHIGALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAV 780  
Qy 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQBELPPPGQEGILLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQBELPPPGQEGILLGFQKAYAPFL 840  
Qy 841 LHWITRGVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900  
Db 841 LHWITRGVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900  
Qy 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFDW 960  
Db 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFDW 960  
Qy 961 LTPSSCCRLYISGPNKDFCPTVNSLNCNCSITMGSVRPSVEQPHKYLPMFLNDRP 1020

[illegible]

Db 1321 ISNFPNNGROF 1332

RESULT 6

US-10-663-208A-4

; Sequence 4, Application US/10663208A

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JB01603K2 US

; CURRENT APPLICATION NUMBER: US/10/663,208A

; CURRENT FILING DATE: 2003-09-16

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17

; PRIOR APPLICATION NUMBER: 10/646,301

; PRIOR FILING DATE: 2003-08-22

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1332

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-663-208A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLGRWLLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

Db 1 MAEAGLGRWLLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLLLOKICPRLVTGNTQACSAKOLVSEASLSITKALLTRCPACSNF 120

Db 61 TPARKITGDHLLLOKICPRLVTGNTQACSAKOLVSEASLSITKALLTRCPACSNF 120

QY 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVAYAEFYQHSFAEQSYDSCSRVRVPA 180

Db 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVAYAEFYQHSFAEQSYDSCSRVRVPA 180

QY 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEGQAVGSGIQPLNEGV 240

Db 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEGQAVGSGIQPLNEGV 240

QY 241 ARCNEQSGDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILICSVPAVVTI 300

Db 241 ARCNEQSGDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILICSVPAVVTI 300

QY 301 LLAGFRVAPARDKSNWDPKKGTSLSKLSFSTHTLLGQFFQGWGTWVASMPLTILVLSV 360

Db 301 LLAGFRVAPARDKSNWDPKKGTSLSKLSFSTHTLLGQFFQGWGTWVASMPLTILVLSV 360

QY 361 IPVVALAAGLVFETLTDPVVELWSPNSQARSEKAFHDOHFGPFPFRTNOVILTAPNRSY 420

Db 361 IPVVALAAGLVFETLTDPVVELWSPNSQARSEKAFHDOHFGPFPFRTNOVILTAPNRSY 420

QY 421 RYDLSLLGPKNFGSLDLDLLELLELQBRHLRHLQVWSPQARNISLQDICYAPLNPDNT 480

Db 421 RYDLSLLGPKNFGSLDLDLLELLELQBRHLRHLQVWSPQARNISLQDICYAPLNPDNT 480

QY 481 SLYDCINSLLOYFQNNRTHLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540

Db 481 SLYDCINSLLOYFQNNRTHLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540

QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAQAKLWEAEFLSEEM 600

Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAQAKLWEAEFLSEEM 600

QY 601 RAFORRNAGMFQVTTFTAEERSLEDEINRTTAEDLPIFATSYVIFLYTSLALGSYSSRSR 660

Db 601 RAFORRNAGMFQVTTFTAEERSLEDEINRTTAEDLPIFATSYVIFLYTSLALGSYSSRSR 660

QY 661 MVDKSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720

Db 661 MVDKSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720

QY 721 YQRLPRRPGPREVHIHGRALGRVAPSMMLCSLSBAICFFLGAITPMPAVRTFALTSLGLAV 780

Db 721 YQRLPRRPGPREVHIHGRALGRVAPSMMLCSLSBAICFFLGAITPMPAVRTFALTSLGLAV 780

QY 781 ILDELLQMSAFVALLSLDSKROEASRLDVCCKVQPOELPRPGQEGELIGFFQKAYAPFL 840

Db 781 ILDELLQMSAFVALLSLDSKROEASRLDVCCKVQPOELPRPGQEGELIGFFQKAYAPFL 840

QY 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

Db 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTTGLGNFSSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLAIPASSWVDDFDW 960

Db 901 YFVTTGLGNFSSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLAIPASSWVDDFDW 960

QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKCMISITMGSVRPSVEQFHKYLPWFLNDRP 1020

Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKCMISITMGSVRPSVEQFHKYLPWFLNDRP 1020

QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALRAARELANITADL 1080

Db 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALRAARELANITADL 1080

QY 1081 RKVPGTDPAPVEPPYTTINVFYQYLTILPEGLPMLSLCLVPTFAVSCILLGLDLRSGLL 1140

Db 1081 RKVPGTDPAPVEPPYTTINVFYQYLTILPEGLPMLSLCLVPTFAVSCILLGLDLRSGLL 1140

QY 1141 NLLSIVMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200

Db 1141 NLLSIVMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200

QY 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQIQLIFFRLNLLITLLGLLHGLVFLPV 1260

Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQIQLIFFRLNLLITLLGLLHGLVFLPV 1260

QY 1261 ILSVVGVDVNPALALEOKRAEBAVAAVMVASCPNHPRSVSTADNIYNNHSEGSIKGAGA 1320

Db 1261 ILSVVGVDVNPALALEOKRAEBAVAAVMVASCPNHPRSVSTADNIYNNHSEGSIKGAGA 1320

QY 1321 ISNFPNNGROF 1332

Db 1321 ISNFPNNGROF 1332

RESULT 7

US-10-736-769-4

; Sequence 4, Application US/10736769

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JB01603-K3-US

; CURRENT APPLICATION NUMBER: US/10/736,769

; CURRENT FILING DATE: 2003-12-16

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17

; PRIOR APPLICATION NUMBER: 10/646,301

; PRIOR FILING DATE: 2003-08-22

; PRIOR APPLICATION NUMBER: 10/663,208

; PRIOR FILING DATE: 2003-09-16

; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-736-769-4

Query Match 100.0%; Score 6909; DB 37; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPTTTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60  
DB 1 MAEAGLRGWLWALLRLAQSEPTTTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSEASISITKALLTRCPACSDNF 120  
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSEASISITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAA 180  
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAA 180

QY 181 ATLAGVTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240  
DB 181 ATLAGVTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240

QY 241 ARCNEGODDVATCSCQDCAASCPAIARPOLDSTFYLGMPSGLVLIILCSVFVAVTI 300  
DB 241 ARCNEGODDVATCSCQDCAASCPAIARPOLDSTFYLGMPSGLVLIILCSVFVAVTI 300

QY 301 LLVGFVRVAPARDKSMVDPKGTSLSDKLSFSTHTLLAGQFFQGGTWTWASWPLTILVLSV 360  
DB 301 LLVGFVRVAPARDKSMVDPKGTSLSDKLSFSTHTLLAGQFFQGGTWTWASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTIDPVELNSAPNSQARSEKAFHDQHPFPFRNQVILTAPRNSY 420  
DB 361 IPVVALAAGLVFTELTIDPVELNSAPNSQARSEKAFHDQHPFPFRNQVILTAPRNSY 420

QY 421 RYDSLILGPKNFSGILDLDLLELLEQLERHLQVWSPQARNISLQDICYAPLAPDNT 480  
DB 421 RYDSLILGPKNFSGILDLDLLELLEQLERHLQVWSPQARNISLQDICYAPLAPDNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTFKDGATLAL 540  
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTFKDGATLAL 540

QY 541 SCHADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600  
DB 541 SCHADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600

QY 601 RAFORMAGMFOVFTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSYSRSV 660  
DB 601 RAFORMAGMFOVFTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSYSRSV 660

QY 661 MVDKATLGLGGVAVVLGVAWMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIVLE 720  
DB 661 MVDKATLGLGGVAVVLGVAWMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIVLE 720

QY 721 YQRLPRRPGPRVHIGRALGRVAPSWMLLCSLSEACFFLGLATPMPAVRTFALTSGLAV 780  
DB 721 YQRLPRRPGPRVHIGRALGRVAPSWMLLCSLSEACFFLGLATPMPAVRTFALTSGLAV 780

QY 781 ILDFLOMSAFVALLSLSKROEASRLDVCCCKVQBELPPGQEGELLIGFPQKAYAPFL 840  
DB 781 ILDFLOMSAFVALLSLSKROEASRLDVCCCKVQBELPPGQEGELLIGFPQKAYAPFL 840

QY 841 LHWITRGVLLLFALFGVSLYSNMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
DB 841 LHWITRGVLLLFALFGVSLYSNMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTTLTGYNFSSEAGNNAICSSAGCANNFSFTQKIYATEPPEQSYLAIPASSWVDDFDW 960

DB 901 YFVTTLTGYNFSSEAGNNAICSSAGCANNFSFTQKIYATEPPEQSYLAIPASSWVDDFDW 960

QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCUKNCSITMGSVRPSVEQFHKYLPWFNDLRP 1020  
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCUKNCSITMGSVRPSVEQFHKYLPWFNDLRP 1020

QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALAAARELANITADL 1080  
DB 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALAAARELANITADL 1080

QY 1081 RKVPGTDPAPFVFPYITITNVFYEQYLTILPEGLFMLSCLCLVPTFAVSCLLGLDLRSGLL 1140  
DB 1081 RKVPGTDPAPFVFPYITITNVFYEQYLTILPEGLFMLSCLCLVPTFAVSCLLGLDLRSGLL 1140

QY 1141 NLLSIVMLVDVTGFMALMDISYNAVSLINLVSAGMSVEFVSHITRSPAIKPTWLER 1200  
DB 1141 NLLSIVMLVDVTGFMALMDISYNAVSLINLVSAGMSVEFVSHITRSPAIKPTWLER 1200

QY 1201 AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLLGLHLGLVLPV 1260  
DB 1201 AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLLGLHLGLVLPV 1260

QY 1261 ILSYVGPDPNPALALEQKRAEEAAMVWASCPNHFSRVSTADNIYVNHSEFSGIKAGA 1320  
DB 1261 ILSYVGPDPNPALALEQKRAEEAAMVWASCPNHFSRVSTADNIYVNHSEFSGIKAGA 1320

QY 1321 ISNFLPNNGROF 1332  
DB 1321 ISNFLPNNGROF 1332

RESULT 8  
US-10-750-386-4  
; Sequence 4: Application US/10750386  
; GENERAL INFORMATION:  
; APPLICANT: Garcia-Calvo, Margarita  
; APPLICANT: Chapman, Kevin  
; APPLICANT: Goulet, Mark  
; APPLICANT: Ujjainwalla, Peroze  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Davis, Chip  
; APPLICANT: Bull, Herb  
; APPLICANT: Thornberry, Nancy A  
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: A36104 074669.0103  
; CURRENT APPLICATION NUMBER: US/10750,386  
; CURRENT FILING DATE: 2003-12-31  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-750-386-4

Query Match 100.0%; Score 6909; DB 37; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPTTTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60  
DB 1 MAEAGLRGWLWALLRLAQSEPTTTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSEASISITKALLTRCPACSDNF 120  
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSEASISITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAA 180  
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAA 180

QY 181 ATLAGVTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240

Db 181 ATLAAGTGCYVGSALCNAQRLWFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEGV 240  
Qy 241 ARCNEQDDVATSCQDCAACPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTI 300  
Db 241 ARCNEQDDVATSCQDCAACPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTI 300  
Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFQCGWGTWASWPLTILVLSV 360  
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFQCGWGTWASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVFTLTDPVELSNAPNSQARSEKAFHQHFGPFRNTQVILTPNRSY 420  
Db 361 IPVVALAAGLVFTLTDPVELSNAPNSQARSEKAFHQHFGPFRNTQVILTPNRSY 420  
Qy 421 RYDSLLGPKNFGSGLDLDLLELELERLRLHQLVMSPEAQRNLSLODICYAPLNPNT 480  
Db 421 RYDSLLGPKNFGSGLDLDLLELELERLRLHQLVMSPEAQRNLSLODICYAPLNPNT 480  
Qy 481 SLYDCCINSLLQYFONNETLALLTANOTLMGOTSQVDMKDFLYCANAPLTFKDGTAAL 540  
Db 481 SLYDCCINSLLQYFONNETLALLTANOTLMGOTSQVDMKDFLYCANAPLTFKDGTAAL 540  
Qy 541 SCMDYGAAPVPFFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600  
Db 541 SCMDYGAAPVPFFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600  
Qy 601 RAPQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWSRV 660  
Db 601 RAPQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWSRV 660  
Qy 661 MVDSKATGLGGAIVVLGAVMAAGFFSVGLTRSSLVLQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDSKATGLGGAIVVLGAVMAAGFFSVGLTRSSLVLQVVPFLVLSVGADNIFIFVLE 720  
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTSLGLAV 780  
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTSLGLAV 780  
Qy 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPQGGEGLLLGFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPQGGEGLLLGFQKAYAPFL 840  
Qy 841 LHWITRGVLLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGPV 900  
Db 841 LHWITRGVLLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGPV 900  
Qy 901 YFVTTILGYNFSEAGMNAICSSAGCNCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNFSEAGMNAICSSAGCNCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960  
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNMSITWGSVRPSVEQPHKYLFWFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNMSITWGSVRPSVEQPHKYLFWFLNDRP 1020  
Qy 1021 NIKCPKGGIAAYSTSVNLTSDQVLAASFMAVHKPLKNSQDYTEALRAARELAANITADL 1080  
Db 1021 NIKCPKGGIAAYSTSVNLTSDQVLAASFMAVHKPLKNSQDYTEALRAARELAANITADL 1080  
Qy 1081 RKVPGTDPAFEPFYTTINVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLSGLL 1140  
Db 1081 RKVPGTDPAFEPFYTTINVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLSGLL 1140  
Qy 1141 NLLSVIMLVDVTGFMALWDISYNAVSLINLVSAVMSVEFVSHITRSPAISTKPTWLER 1200  
Db 1141 NLLSVIMLVDVTGFMALWDISYNAVSLINLVSAVMSVEFVSHITRSPAISTKPTWLER 1200  
Qy 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLQIIPFFRLNLLITLLGLHLGVFLPV 1260  
Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLQIIPFFRLNLLITLLGLHLGVFLPV 1260  
Qy 1261 ILSYVGPDPVNPALAEQKRAEBAVAAMVVASCPNHPRSRVSTADNIYVNHSEFGSIKGAGA 1320

1261 ILSYVGPDPVNPALAEQKRAEBAVAAMVVASCPNHPRSRVSTADNIYVNHSEFGSIKGAGA 1320  
Qy 1321 ISNFLPNNGRQF 1332  
Db 1321 ISNFLPNNGRQF 1332  
RESULT 9  
PCT-US01-04098A-1831  
; Sequence 1831, Application PC/TUS0104098A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-029  
; CURRENT APPLICATION NUMBER: PCT/US01/04098A  
; CURRENT FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: Not Yet Assigned  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/728,422  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 09/693,325  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/663,561  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/654,936  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/620,325  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/598,075  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 3960  
; SOFTWARE: Custom  
; SEQ ID NO 1831  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-04098A-1831  
Query Match 99.8%; Score 6896; DB 1; Length 1332;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MAEAGLRGWLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGLMTLSNVCLSN 60  
Db 1 MAEAGLRGWLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGLMTLSNVCLSN 60  
Qy 61 TPARKITGDHLLILQKICPLRYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILQKICPLRYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120  
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDCSRVRVPA 180  
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDCSRVRVPA 180  
Qy 181 ATLAAGTGCYVGSALCNAQRLWFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEGV 240  
Db 181 ATLAAGTGCYVGSALCNAQRLWFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEGV 240  
Qy 241 ARCNEQDDVATSCQDCAACPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTI 300  
Db 241 ARCNEQDDVATSCQDCAACPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTI 300  
Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFQCGWGTWASWPLTILVLSV 360  
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFQCGWGTWASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVFTLTDPVELSNAPNSQARSEKAFHQHFGPFRNTQVILTPNRSY 420  
Db 361 IPVVALAAGLVFTLTDPVELSNAPNSQARSEKAFHQHFGPFRNTQVILTPNRSY 420

QY 421 RYDSSLGPKNFSGIILDLDDLLLELLELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNT 480  
DB 421 RYDSSLGPKNFSGIILDLDDLLLELLELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGATLAL 540  
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGATLAL 540  
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
QY 601 RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
DB 601 RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
QY 661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720  
DB 661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720  
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTMPAVRTFALTSGLAV 780  
DB 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTMPAVRTFALTSGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQPELPPPGQEGELLGFFQKAYAPFL 840  
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQPELPPPGQEGELLGFFQKAYAPFL 840  
QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDFLFLNRYFEVGAPV 900  
DB 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDFLFLNRYFEVGAPV 900  
QY 901 YFVTTLGYNFSSAGNNAICSSAGCNFFTKIQYATEFPPEQSYLAIIPASSWDDFIDW 960  
DB 901 YFVTTLGYNFSSAGNNAICSSAGCNFFTKIQYATEFPPEQSYLAIIPASSWDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLNKMCSITMGSVRPSVEQFHKILPWFLNDRP 1020  
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLNKMCSITMGSVRPSVEQFHKILPWFLNDRP 1020  
QY 1021 NIKKCPKGLAAYSTSVNLTSDQVTLASRFMAHKLKNSQDTEALRAARELAANITADL 1080  
DB 1021 NIKKCPKGLAAYSTSVNLTSDQVTLASRFMAHKLKNSQDTEALRAARELAANITADL 1080  
QY 1081 RKVPGTDPAFEVFPYITINVFYEOYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
DB 1081 RKVPGTDPAFEVFPYITINVFYEOYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
QY 1141 NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200  
DB 1141 NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200  
QY 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLQIIFPFRNLNLTLLGLLHGLVFLPV 1260  
DB 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLQIIFPFRNLNLTLLGLLHGLVFLPV 1260  
QY 1261 ILSYVGPDVNPALAEOKRAEBAVAAMVVASCPNHPRSRVSTADNIVVNHSPGSKGAGA 1320  
DB 1261 ILSYVGPDVNPALAEOKRAEBAVAAMVVASCPNHPRSRVSTADNIVVNHSPGSKGAGA 1320  
QY 1321 ISNFLPNNGRQF 1332  
DB 1321 ISNFLPNNGRQF 1332

RESULT 10

US-10-239-316-9  
; Sequence 9, Application US/10239316  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: KITA, Shunbun  
; APPLICANT: SATOMI, Tomoko Komiyama

; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/10/239,316  
; CURRENT FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 9  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-239-316-9

Query Match 99.8%; Score 6896; DB 32; Length 1332;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKPELSGSLMTLSNVCSLSN 60  
DB 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKPELSGSLMTLSNVCSLSN 60  
QY 61 TPARKITGDHLLILOKICPRLYTGNPTQACCSAKQLVSLSEASITTKALLTRCPACSDNF 120  
DB 61 TPARKITGDHLLILOKICPRLYTGNPTQACCSAKQLVSLSEASITTKALLTRCPACSDNF 120  
QY 121 VNLHCHTCSNPSLFINVTRVAQAGOLPAVVAEYQHSFAESQSDSCSRVRVPA 180  
DB 121 VNLHCHTCSNPSLFINVTRVAQAGOLPAVVAEYQHSFAESQSDSCSRVRVPA 180  
QY 181 ATLAVGTMCVYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIQPLNEGV 240  
DB 181 ATLAVGTMCVYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIQPLNEGV 240  
QY 241 ARCNSQGGDVATCSCQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300  
DB 241 ARCNSQGGDVATCSCQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300  
QY 301 LLVGRFVAPARDKSKWDPKKTGTSLSDKLSFSTHTLLGQFQGWGTWASWPLTILVLSV 360  
DB 301 LLVGRFVAPARDKSKWDPKKTGTSLSDKLSFSTHTLLGQFQGWGTWASWPLTILVLSV 360  
QY 361 IPVVALAAGLVFETLTTPVELWSAPNSQARSEKAFHDOHFGFFFRNTNOVILTAPNRSSY 420  
DB 361 IPVVALAAGLVFETLTTPVELWSAPNSQARSEKAFHDOHFGFFFRNTNOVILTAPNRSSY 420  
QY 421 RYDSSLGPKNFSGIILDLDDLLLELLELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNT 480  
DB 421 RYDSSLGPKNFSGIILDLDDLLLELLELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGATLAL 540  
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGATLAL 540  
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
QY 601 RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
DB 601 RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
QY 661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720  
DB 661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720  
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTMPAVRTFALTSGLAV 780  
DB 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTMPAVRTFALTSGLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQPELPPPGQEGELLGFFQKAYAPFL 840  
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKQPELPPPGQEGELLGFFQKAYAPFL 840

Db 781 ILDFLLQMSAFVALLSLDSKQEARSLDVCCKVPQELPPPGQEGILLGLFFQKAYAPFL 840  
Qy 841 LHWITRGVLLLFALFGVSLXSNCHISVGLDDELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHWITRGVLLLFALFGVSLXSNCHISVGLDDELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSTOKIOVATEPEOSYLAIPASSWVDDPFDW 960  
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSTOKIOVATEPEOSYLAIPASSWVDDPFDW 960  
Qy 961 LTPSSCCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRPSVEQFHXYLPWFNDRP 1020  
Db 961 LTPSSCCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRPSVEQFHXYLPWFNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Qy 1081 RKVPGTDPAFVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
Db 1081 RKVPGTDPAFVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
Qy 1141 NLLSIVMLVDVTFGMALWDISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLER 1200  
Db 1141 NLLSIVMLVDVTFGMALWGISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLER 1200  
Qy 1201 AKEATISMGSAVFAVAGVAMTNLPGILVLGLAKAQIQTFFRNLNLTLLGLLHGLVLPV 1260  
Db 1201 AKEATISMGSAVFAVAGVAMTNLPGILVLGLAKAQIQTFFRNLNLTLLGLLHGLVLPV 1260  
Qy 1261 ILSYVGPDPNPALALEQKRAEAAVAVMVASCPNHPBSRVSTADNIIYNHSPFGSIKGAGA 1320  
Db 1261 ILSYVGPDPNPALALEQKRAEAAVAVMVASCPNHPBSRVSTADNIIYNHSPFGSIKGAGA 1320  
Qy 1321 ISNFLPNNGROF 1332  
Db 1321 ISNFLPNNGROF 1332

RESULT 11

US-10-258-899A-1831  
; Sequence 1831, Application US/10258899A  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Drmanac, Radoje T.  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhou, Ping  
; APPLICANT: Xu, Chongjun  
; APPLICANT: Cao, Yicheng  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhi Wei  
; APPLICANT: Xue, Aidong  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Goodrich, Kyle  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 787CIP2-2G/US  
; CURRENT APPLICATION NUMBER: US/10/258,899A  
; CURRENT FILING DATE: 2003-11-06  
; PRIOR APPLICATION NUMBER: PCT/US01/04098  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 09/774,434  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/728,422  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 09/693,325

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/663,561  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/654,936  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/620,325  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/598,075  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 3960  
; SOFTWARE: Custom  
; SEQ ID NO 1831  
; TYPE: PRT  
; LENGTH: 1332  
; ORGANISM: Homo sapiens  
; US-10-258-899A-1831  
  
Query Match 99.8%; Score 6896; DB 32; Length 1332;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 MAEAGLRGMLLWALLRLAQSEPPYTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60  
Db 1 MAEAGLRGMLLWALLRLAQSEPPYTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60  
Qy 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120  
Qy 121 VNLHCHNTCSPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFAQSVDSCSRVRPAA 180  
Db 121 VNLHCHNTCSPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFAQSVDSCSRVRPAA 180  
Qy 181 ATLAVGTMCGVYGSALCNQRLNFGQDTGNGLAPLDITPHLLEPGQAVSGIQPLNEGV 240  
Db 181 ATLAVGTMCGVYGSALCNQRLNFGQDTGNGLAPLDITPHLLEPGQAVSGIQPLNEGV 240  
Qy 241 ARCNESQDDVATCSCQDCAASCPATARPQALDSTFVILGMPGSLVLIILCSVFAVVTI 300  
Db 241 ARCNESQDDVATCSCQDCAASCPATARPQALDSTFVILGMPGSLVLIILCSVFAVVTI 300  
Qy 301 LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGMTWVASWPLTILVLSV 360  
Db 301 LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGMTWVASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSSY 420  
Db 361 IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSSY 420  
Qy 421 RYDSLLGPKNFGSILDLLELLELOERLHLQVWSPQAQRNLSIQDICYAPLNPDNT 480  
Db 421 RYDSLLGPKNFGSILDLLELLELOERLHLQVWSPQAQRNLSIQDICYAPLNPDNT 480  
Qy 481 SLXDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTAAL 540  
Db 481 SLXDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTAAL 540  
Qy 541 SCMDYGAAPVFPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEEM 600  
Db 541 SCMDYGAAPVFPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEEM 600  
Qy 601 RAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSSSRSR 660  
Db 601 RAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSSSRSR 660  
Qy 661 MVDKATLGLGVAVVLGAVNMAAMGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGVAVVLGAVNMAAMGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVLE 720

QY 721 YORLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSGSLAV 780  
Db 721 YORLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSGSLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGGLLLGFFOKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGGLLLGFFOKAYAPFL 840  
QY 841 LHMITRGVULLFLALFGVSLYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHMITRGVULLFLALFGVSLYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTILGYNPFSSAGNNAICSSAGCANNFSFTQIKIYATEPPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNPFSSAGNNAICSSAGCANNFSFTQIKIYATEPPEQSYLAIPASSWVDDFIDW 960  
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNLKNCMSITMGSVRPSVQFHKYLPWFLENDRP 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNLKNCMSITMGSVRPSVQFHKYLPWFLENDRP 1020  
QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080  
Db 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080  
QY 1081 RKVPGTDPAPFEPVPIYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
Db 1081 RKVPGTDPAPFEPVPIYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140  
QY 1141 NLLSIVMILVDTGFWALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISKPTWLER 1200  
Db 1141 NLLSIVMILVDTGFWALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISKPTWLER 1200  
QY 1201 AKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLLTLGLLHGLVFLPV 1260  
Db 1201 AKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLLTLGLLHGLVFLPV 1260  
QY 1261 ILSYVGPDPNPAALQKRAEAAVAAVMVASCPNHPSPRSTADNIYVNSHSPGSIKAGA 1320  
Db 1261 ILSYVGPDPNPAALQKRAEAAVAAVMVASCPNHPSPRSTADNIYVNSHSPGSIKAGA 1320  
QY 1321 ISNFLPNNGROP 1332  
Db 1321 ISNFLPNNGROP 1332

RESULT 12  
US-10-293-244-1831  
; Sequence 1831, Application US/10293244  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-029  
; CURRENT APPLICATION NUMBER: US/10/293,244  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: Not Yet Assigned  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/728,422  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 09/693,325  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/663,561  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/654,936  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/620,325  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/598,075  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03

; NUMBER OF SEQ ID NOS: 3960  
; SOFTWARE: Custom  
; SEQ ID NO 1831  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-244-1831

Query Match 99.8%; Score 6896; DB 32; Length 1332;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGURGMLLWALLRLAQSEBYTTIHQPGYCAFYDECKNPDELSGSLMTLSNVCSLSN 60  
Db 1 MAEAGURGMLLWALLRLAQSEBYTTIHQPGYCAFYDECKNPDELSGSLMTLSNVCSLSN 60  
QY 61 TPARKITGDHLILLQKICPRLYTGNPTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLILLQKICPRLYTGNPTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120  
QY 121 VNLHCHNTCSPNOSLFINVTRVAQLCAGQLPAVVAEAFYQHSFAEQSDSCSRVRVPAA 180  
Db 121 VNLHCHNTCSPNOSLFINVTRVAQLCAGQLPAVVAEAFYQHSFAEQSDSCSRVRVPAA 180  
QY 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240  
Db 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240  
QY 241 ARCNSQGGDDVATCSCQDCAASPATARPQALDSTPYLGQMPGSLVLIILCSVPFVAVTI 300  
Db 241 ARCNSQGGDDVATCSCQDCAASPATARPQALDSTPYLGQMPGSLVLIILCSVPFVAVTI 300  
QY 301 LLVGRFVAPARDKSKMVDPKGTSLSDKLSFSFTHTLGGFCGSGTGWASWPLTILVLSV 360  
Db 301 LLVGRFVAPARDKSKMVDPKGTSLSDKLSFSFTHTLGGFCGSGTGWASWPLTILVLSV 360  
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQFHFRTNQVILTAPNRSSY 420  
Db 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQFHFRTNQVILTAPNRSSY 420  
QY 421 RYDSLLGPKNFGSGLDLDLLELELQERLRLHQLWSPQASRNISLQDICVAPLNPONT 480  
Db 421 RYDSLLGPKNFGSGLDLDLLELELQERLRLHQLWSPQASRNISLQDICVAPLNPONT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540  
Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540  
QY 541 SCWADYGAPVFPFPLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
Db 541 SCWADYGAPVFPFPLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600  
QY 601 RAFQRRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISALGSYSWSRV 660  
Db 601 RAFQRRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISALGSYSWSRV 660  
QY 661 MYDSKATIGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGAADNIFIVLE 720  
Db 661 MYDSKATIGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGAADNIFIVLE 720  
QY 721 YORLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSGSLAV 780  
Db 721 YORLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSGSLAV 780  
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGGLLLGFFOKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQEGGLLLGFFOKAYAPFL 840  
QY 841 LHMITRGVULLFLALFGVSLYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHMITRGVULLFLALFGVSLYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTILGYNPFSSAGNNAICSSAGCANNFSFTQIKIYATEPPEQSYLAIPASSWVDDFIDW 960

Db 901 YFVTTLGYNFSSEAGWNAICSSAGCNFSFTQKIYATEPEQSYLAI PASSWDDFDW 960  
Qy 961 LTPSSCCRLYISGNKDKPCPSTVNSLNCNKMSITMGSVRPSVEQFKYLPWFLNDRP 1020  
Db 961 LTPSSCCRLYISGNKDKPCPSTVNSLNCNKMSITMGSVRPSVEQFKYLPWFLNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFVAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFVAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Qy 1081 RKVPGTDPAFEVFPYTTTNVFEYQYLITLPEGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140  
Db 1081 RKVPGTDPAFEVFPYTTTNVFEYQYLITLPEGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140  
Qy 1141 NLSIVMLVDTVGFMAIWDISYNAVSLINLVSAGMSVEFVSHITSPFAISTKPTWLER 1200  
Db 1141 NLSIVMLVDTVGFMAIWDISYNAVSLINLVSAGMSVEFVSHITSPFAISTKPTWLER 1200  
Qy 1201 AKEATISNGSAVFAVAMTNLPGILVLGLAKAQLIQIFFRLNLITLLGLLHGLVFLPV 1260  
Db 1201 AKEATISNGSAVFAVAMTNLPGILVLGLAKAQLIQIFFRLNLITLLGLLHGLVFLPV 1260  
Qy 1261 ILSYVGPDVNPALAEQKRAEEAAVAVVAVASCPNHPKSVSTADNIYVNHSPGSIKAGA 1320  
Db 1261 ILSYVGPDVNPALAEQKRAEEAAVAVVAVASCPNHPKSVSTADNIYVNHSPGSIKAGA 1320  
Qy 1321 ISNPLPNNGRQF 1332  
Db 1321 ISNPLPNNGRQF 1332

RESULT 13

US-11-218-141-1831  
; Sequence 1831, Application US/11218141  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-029  
; CURRENT APPLICATION NUMBER: US/11/218,141  
; CURRENT FILING DATE: 2005-08-31  
; PRIOR APPLICATION NUMBER: Not Yet Assigned  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/728,422  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 09/693,325  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/663,561  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/654,936  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/620,325  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/598,075  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 3960  
; SOFTWARE: Custom  
; SEQ ID NO 1831  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-218-141-1831

Query Match 99.8%; Score 6896; DB 42; Length 1332;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60

Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60  
Qy 61 TPARKITGDHLLILLOKICPRLYTGNPTQACCSAKOLVLEASISITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILLOKICPRLYTGNPTQACCSAKOLVLEASISITKALLTRCPACSDNF 120  
Qy 121 VNLHCHTNCSPNOSLFINTVTRVAQLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180  
Db 121 VNLHCHTNCSPNOSLFINTVTRVAQLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180  
Qy 181 ATLAVGTCMGVYSALCNAQRWLNFGQDGTGNGLAPLDITPHLLRPGQAVSGSQPLNEGV 240  
Db 181 ATLAVGTCMGVYSALCNAQRWLNFGQDGTGNGLAPLDITPHLLRPGQAVSGSQPLNEGV 240  
Qy 241 ARCNESQGDVAVTCQDCAASCPAIPALPOLDSTFYLQMPGSLVLIILCSVFAVVTI 300  
Db 241 ARCNESQGDVAVTCQDCAASCPAIPALPOLDSTFYLQMPGSLVLIILCSVFAVVTI 300  
Qy 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWPLTILVLSV 360  
Db 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVPTTELTTDPVELWSAPNSQARSSEKAFHDQHFQFFRTNQVILTAPNRSY 420  
Db 361 IPVVALAAGLVPTTELTTDPVELWSAPNSQARSSEKAFHDQHFQFFRTNQVILTAPNRSY 420  
Qy 421 RYDSLLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 480  
Db 421 RYDSLLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480  
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540  
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540  
Qy 541 SCWADYGAPVFPFLAIGGYKDYSEABALIMTFSLNNYPAGDRLAQAALWEAFLEEM 600  
Db 541 SCWADYGAPVFPFLAIGGYKDYSEABALIMTFSLNNYPAGDRLAQAALWEAFLEEM 600  
Qy 601 RAFQRRMAGMFQVTFMAERSLEDEINRTAEDLPIPATSVIVIFLYISLALGSYSSSRV 660  
Db 601 RAFQRRMAGMFQVTFMAERSLEDEINRTAEDLPIPATSVIVIFLYISLALGSYSSSRV 660  
Qy 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSVLQVVPFLVSVGADNIFIFVLE 720  
Db 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSVLQVVPFLVSVGADNIFIFVLE 720  
Qy 721 YQRLPRPGEPREHIGRALGRVAPSMILCSLSEAI CFFILGALTPMBAVTFALTSGLV 780  
Db 721 YQRLPRPGEPREHIGRALGRVAPSMILCSLSEAI CFFILGALTPMBAVTFALTSGLV 780  
Qy 781 ILDFLLQMSAFVALLSLDSKQREASRLDCCVKPQELPPPGQEGLLLGFFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKQREASRLDCCVKPQELPPPGQEGLLLGFFQKAYAPFL 840  
Qy 841 LHMWITRGVLLLFALFGVSLYSMCHISVGLDQBELPKDSYLLDYLFLNRYFEVGA 900  
Db 841 LHMWITRGVLLLFALFGVSLYSMCHISVGLDQBELPKDSYLLDYLFLNRYFEVGA 900  
Qy 901 YFVTTLGYNFSSEAGWNAICSSAGCNFSFTQKIYATEPEQSYLAI PASSWDDFDW 960  
Db 901 YFVTTLGYNFSSEAGWNAICSSAGCNFSFTQKIYATEPEQSYLAI PASSWDDFDW 960  
Qy 961 LTPSSCCRLYISGNKDKPCPSTVNSLNCNKMSITMGSVRPSVEQFKYLPWFLNDRP 1020  
Db 961 LTPSSCCRLYISGNKDKPCPSTVNSLNCNKMSITMGSVRPSVEQFKYLPWFLNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFVAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFVAYHKKPLKNSQDYTEALRAARELANITADL 1080  
Qy 1081 RKVPGTDPAFEVFPYTTTNVFEYQYLITLPEGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140

Db 1081 RKVPGTDPAFEPVPTITNVFYEQYLITLPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140  
QY 1141 NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAVGMSVEFVSHITRSPAISTKPTWLER 1200  
Db 1141 NLLSIVMILVDTVGFMAWLGISTYNAVSLINLVSAVGMSVEFVSHITRSPAISTKPTWLER 1200  
QY 1201 AKEATISMGSAVPAVAVMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLLGLHLGLVPLPV 1260  
Db 1201 AKEATISMGSAVPAVAVMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLLGLHLGLVPLPV 1260  
QY 1261 ILSYVGPDPNPALALEOKRAEEAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGSIKAGA 1320  
Db 1261 ILSYVGPDPNPALALEOKRAEEAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGSIKAGA 1320  
QY 1321 ISNPLPNNGRQF 1332  
Db 1321 ISNPLPNNGRQF 1332

RESULT 14  
US-11-242-459-9  
; Sequence 9, Application US/11242459  
; GENERAL INFORMATION:  
; APPLICANT: TANIYAMA, Yoshio  
; APPLICANT: KITA, Shunbun  
; APPLICANT: SATOMI, Tomoko Komiyama  
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof  
; FILE REFERENCE: 2703USOP  
; CURRENT APPLICATION NUMBER: US/11/242,459  
; CURRENT FILING DATE: 2005-10-03  
; PRIOR APPLICATION NUMBER: US/10/239,316  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: PCT/JP01/02279  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: JP2000-088595  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 59  
; SEQ ID NO 9  
; LENGTH: 1332  
; TYPE: PRT  
; ORGANISM: Human  
US-11-242-459-9

Query Match 99.8%; Score 6896; DB 42; Length 1332;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN 60  
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN 60  
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120  
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120  
QY 121 VNLHCNTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAPQSYDSCSRVRVPAA 180  
Db 121 VNLHCNTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAPQSYDSCSRVRVPAA 180  
QY 181 ATLAVGTCGVYGSALCNAQRMWLNFOGDTGNGLAPLIDITFHLLEPQAVGSGIQPLNEGV 240  
Db 181 ATLAVGTCGVYGSALCNAQRMWLNFOGDTGNGLAPLIDITFHLLEPQAVGSGIQPLNEGV 240  
QY 241 ARCNEQSGDDVATCSCQDCCAAPARPAQALDSTFVLGOMPGLVLIILCSVFAVVTI 300  
Db 241 ARCNEQSGDDVATCSCQDCCAAPARPAQALDSTFVLGOMPGLVLIILCSVFAVVTI 300  
QY 301 LLVGFVRVAPARSKMVDPKKGTSLSDKLSFSHTLLGQFGQGTWVASWPLTILVLSV 360  
Db 301 LLVGFVRVAPARSKMVDPKKGTSLSDKLSFSHTLLGQFGQGTWVASWPLTILVLSV 360  
QY 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFDQHFQFPFRTNQVILTAPNRSSY 420

Db 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFDQHFQFPFRTNQVILTAPNRSSY 420  
QY 421 RYDSLLGPKNFGSILDLDDLLLELLELQERLRLQWMSPEAQRNISLQDICYAPLNPONT 480  
Db 421 RYDSLLGPKNFGSILDLDDLLLELLELQERLRLQWMSPEAQRNISLQDICYAPLNPONT 480  
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDMKDHFLYCANAPLTFKDGATLAL 540  
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDMKDHFLYCANAPLTFKDGATLAL 540  
QY 541 SCHADYGAPVFPPLATGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAOKLWBEAFLEEM 600  
Db 541 SCHADYGAPVFPPLATGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAOKLWBEAFLEEM 600  
QY 601 RAFQRMAGMFQVFTTAERSLEDEINRTTAEDLPIFATSYIVFLYISLAGSYSSWSRV 660  
Db 601 RAFQRMAGMFQVFTTAERSLEDEINRTTAEDLPIFATSYIVFLYISLAGSYSSWSRV 660  
QY 661 MVDKATLGLGGVAVVVLGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSVGADNIFIPVLE 720  
Db 661 MVDKATLGLGGVAVVVLGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSVGADNIFIPVLE 720  
QY 721 YQRLPRPGEPREVIHGRALGRVAPSKMLCSLSEATCFELGALTMPAVRTFALTSGIAY 780  
Db 721 YQRLPRPGEPREVIHGRALGRVAPSKMLCSLSEATCFELGALTMPAVRTFALTSGIAY 780  
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGGEGILLGFFOKAVAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCVKPOELPPPGGEGILLGFFOKAVAPFL 840  
QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
Db 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900  
QY 901 YFVTTILGYNFSSEAGNVAICSSAGCNPFSTOKIQVATEFPPEQSYLAIPASSWDDPIDW 960  
Db 901 YFVTTILGYNFSSEAGNVAICSSAGCNPFSTOKIQVATEFPPEQSYLAIPASSWDDPIDW 960  
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMCSITMGSVRPSVSEQFHKLWFLNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMCSITMGSVRPSVSEQFHKLWFLNDRP 1020  
QY 1021 NIKCPKGGLAAYSTSVNLTSDGQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080  
Db 1021 NIKCPKGGLAAYSTSVNLTSDGQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080  
QY 1081 RKVPGTDPAFEPVPTITNVFYEQYLITLPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140  
Db 1081 RKVPGTDPAFEPVPTITNVFYEQYLITLPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140  
QY 1141 NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAVGMSVEFVSHITRSPAISTKPTWLER 1200  
Db 1141 NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAVGMSVEFVSHITRSPAISTKPTWLER 1200  
QY 1201 AKEATISMGSAVPAVAVMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLLGLHLGLVPLPV 1260  
Db 1201 AKEATISMGSAVPAVAVMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLLGLHLGLVPLPV 1260  
QY 1261 ILSYVGPDPNPALALEOKRAEEAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGSIKAGA 1320  
Db 1261 ILSYVGPDPNPALALEOKRAEEAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGSIKAGA 1320  
QY 1321 ISNPLPNNGRQF 1332  
Db 1321 ISNPLPNNGRQF 1332

RESULT 15  
PCT-US01-04098A-1830  
; Sequence 1830, Application PC/TUS0104098A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

Mon Apr 10 07:16:15 2006

FILE REFERENCE: 21272-029  
CURRENT APPLICATION NUMBER: PCT/US01/04098A  
CURRENT FILING DATE: 2001-02-05  
PRIOR APPLICATION NUMBER: Not Yet Assigned  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: 09/728,422  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 09/693,325  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 09/663,561  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 09/654,936  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 09/620,325  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/598,075  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: 09/560,875  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 09/496,914  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 3960  
SOFTWARE: Custom  
SEQ ID NO 1830  
LENGTH: 1359  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-04098A-1830

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLAQAEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLN 60  
Db 1 MAEAGLRGWLWALLRLAQAEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLN 60  
Qy 61 TPARKITGDHLLQKICPRLTYGNTQACSAKQVLSLEASLSITKALLTRCPACSNF 120  
Db 61 TPARKITGDHLLQKICPRLTYGNTQACSAKQVLSLEASLSITKALLTRCPACSNF 120  
Qy 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVAYEAFYQHSFPAEQSDSCSRVRVPA 180  
Db 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVAYEAFYQHSFPAEQSDSCSRVRVPA 180  
Qy 181 ATLAVTGTCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAQVSGIQPLNEGV 240  
Db 181 ATLAVTGTCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAQVSGIQPLNEGV 240  
Qy 241 ARCNEQGDVATCSCQDCAAPAIARPOALDSTFYLGOMPGSLVLIILLCVFAVVTI 300  
Db 241 ARCNEQGDVATCSCQDCAAPAIARPOALDSTFYLGOMPGSLVLIILLCVFAVVTI 300  
Qy 301 LLVGFVAPARDKSKWDPKKGTSLSKLSFSTHLLGQFFQCGWTWASWPLTILVLSV 360  
Db 301 LLVGFVAPARDKSKWDPKKGTSLSKLSFSTHLLGQFFQCGWTWASWPLTILVLSV 360  
Qy 361 IPVVALAAGLVFTLTDPVELMSAPNSQARSEKAFHDQHFQFFFTNQVILTAPNRSY 420  
Db 361 IPVVALAAGLVFTLTDPVELMSAPNSQARSEKAFHDQHFQFFFTNQVILTAPNRSY 420  
Qy 421 RYDSSLGPKNFGSGLDLDLLELLELLELRLHQLQWSPQAORNSISLQDICVAPLNPDNT 480  
Db 421 RYDSSLGPKNFGSGLDLDLLELLELLELRLHQLQWSPQAORNSISLQDICVAPLNPDNT 480  
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVWDKDHFLYCANAPLTFKDGFTALAL 540  
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVWDKDHFLYCANAPLTFKDGFTALAL 540  
Qy 541 SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600  
Db 541 SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600

Qy 601 RAFORMMAGMFOVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
Db 601 RAFORMMAGMFOVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660  
Qy 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVLOVVPFLVLSVGAONIIFIVLE 720  
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVLOVVPFLVLSVGAONIIFIVLE 720  
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSISEAICPFLGALTTPMPAVRTFALTSLAV 780  
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSISEAICPFLGALTTPMPAVRTFALTSLAV 780  
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQDELPPPGQEGELLGFFQKAYAPFL 840  
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQDELPPPGQEGELLGFFQKAYAPFL 840  
Qy 841 LHMWTRGVVLLFLALFGVLSYSMCHIISVGLQDELALPKDSYLLDYFLFLNRFFEVPAPV 900  
Db 841 LHMWTRGVVLLFLALFGVLSYSMCHIISVGLQDELALPKDSYLLDYFLFLNRFFEVPAPV 900  
Qy 901 YFVTTILGYNFSSEAGMNAICSSAGCNMFSTQKIQYATFEPPEQSYLAIPASSWVDDFIDW 960  
Db 901 YFVTTILGYNFSSEAGMNAICSSAGCNMFSTQKIQYATFEPPEQSYLAIPASSWVDDFIDW 960  
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSTMGSVRPSVQFHKYLPWFNDPR 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSTMGSVRPSVQFHKYLPWFNDPR 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSDGOVL-----ASRFMAVH 1053  
Db 1021 NIKCPKGLAAYSTSVNLTSDGOVL-----ASRFMAVH 1053  
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPFPYTTITNVFYEQYLTILPEGL 1113  
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPFPYTTITNVFYEQYLTILPEGL 1113  
Qy 1114 FMLSCLVPTFAVSCLLGLDLRLSGLINLLSIIVMILVDTVGFMAWDISYNAVSLINLVS 1200  
Db 1141 FMLSCLVPTFAVSCLLGLDLRLSGLINLLSIIVMILVDTVGFMAWDISYNAVSLINLVS 1200  
Qy 1174 AVGMSVEFVSHITTESFAISTKPTWLERAKATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITTESFAISTKPTWLERAKATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260  
Qy 1234 LIQIFFFRLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEOKRAEEAAVAAVWVASC 1293  
Db 1261 LIQIFFFRLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEOKRAEEAAVAAVWVASC 1320  
Qy 1294 NHPRSVSTADNIYVNHSPFEGSIKAGAIISNPLNNGRQF 1332  
Db 1321 NHPRSVSTADNIYVNHSPFEGSIKAGAIISNPLNNGRQF 1359

Search completed: April 7, 2006, 19:32:35  
Job time : 576 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:23:20 ; Search time 77 Seconds  
(without alignments)  
2380.515 Million cell updates/sec

Title: US-10-736-769-4  
Perfect score: 6909  
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGISNFIPLNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 594368 seqs, 137612332 residues

Total number of hits satisfying chosen parameters: 594368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New\*

- 1: /SIDS5/ptodata/1/paa/PCT NEW COMB.pcp.\*
- 2: /SIDS5/ptodata/1/paa/US06 NEW COMB.pcp.\*
- 3: /SIDS5/ptodata/1/paa/US07 NEW COMB.pcp.\*
- 4: /SIDS5/ptodata/1/paa/US08 NEW COMB.pcp.\*
- 5: /SIDS5/ptodata/1/paa/US09 NEW COMB.pcp.\*
- 6: /SIDS5/ptodata/1/paa/US10 NEW COMB.pcp.\*
- 7: /SIDS5/ptodata/1/paa/US11 NEW COMB.pcp.\*
- 8: /SIDS5/ptodata/1/paa/US60 NEW COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	6	US-11-301-094-6
2	6872.5	99.5	1359	6	US-11-270-796-22
3	5407	78.3	1333	6	US-11-270-796-3
4	2402.5	34.8	1278	1	PCT-US06-05584-822
5	2402.5	34.8	1278	6	US-11-191-274A-129
6	2402.5	34.8	1278	6	US-11-191-274A-130
7	2402.5	34.8	1278	7	US-11-385-692-2400
8	2402.5	34.8	1278	7	US-11-385-692-2401
9	1341.5	19.4	1363	6	US-11-301-094-2
10	1046	15.1	1274	6	US-11-301-094-4
11	782.5	11.3	419	6	US-10-953-349-21065
12	588.5	8.5	967	6	US-10-461-673-16739
13	583.5	8.4	1182	6	US-11-332-764-2
14	583.5	8.4	1182	6	US-11-337-244-149
15	482	7.0	891	8	US-60-772-265-1197
16	455	6.6	1358	8	US-60-772-265-241
17	437	6.3	831	7	US-10-360-355-133805
18	349	5.1	783	6	US-10-461-673-16750
19	325	4.7	204	7	US-11-360-355-141351
20	261	3.8	575	7	US-11-360-355-120785
21	252	3.6	505	6	US-10-461-673-16871
22	229.5	3.3	455	7	US-11-360-355-141367
23	225	3.3	542	6	US-11-214-063A-1670
24	222.5	3.2	422	7	US-11-360-355-141365
25	219	3.2	1137	6	US-10-461-673-16889

26	219	3.2	1330	6	US-10-461-673-10541	Sequence 10541, A
27	190	2.8	126	7	US-11-360-355-141352	Sequence 141352,
28	152.5	2.2	632	6	US-11-293-697-3548	Sequence 3548, Ap
29	141	2.0	164	7	US-11-360-355-127069	Sequence 127069,
30	139.5	2.0	484	7	US-11-360-355-120169	Sequence 120169,
31	135.5	2.0	252	7	US-11-360-355-140206	Sequence 140206,
32	126	1.8	808	6	US-10-724-972B-7050	Sequence 7050, Ap
33	123.5	1.8	171	7	US-11-360-355-149589	Sequence 149589,
34	121.5	1.8	1043	6	US-10-536-606-20	Sequence 20, Appl
35	120.5	1.7	2426	6	US-11-203-806A-11	Sequence 11, Appl
36	118.5	1.7	788	6	US-11-214-063A-2044	Sequence 2044, Ap
37	118.5	1.7	788	6	US-11-214-063A-1692	Sequence 1692, Ap
38	117.5	1.7	619	8	US-60-732-162-1828	Sequence 1828, Ap
39	116.5	1.7	619	6	US-11-312-958-46	Sequence 46, Appl
40	116.5	1.7	619	6	US-10-461-673-12167	Sequence 12167, A
41	114	1.7	488	8	US-60-752-355-45206	Sequence 45206, A
42	114	1.7	985	6	US-11-293-697-2874	Sequence 2874, Ap
43	112.5	1.6	697	6	US-10-703-799B-226	Sequence 226, App
44	111	1.6	201	7	US-11-360-355-152528	Sequence 152528,
45	111	1.6	958	8	US-60-781-953-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-11-301-094-6

- ; Sequence 6, Application US/11301094
- ; GENERAL INFORMATION:
- ; APPLICANT: Levitan, Diane J
- ; APPLICANT: Smith, Marsha
- ; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
- ; FILE REFERENCE: JB06242US01
- ; CURRENT APPLICATION NUMBER: US/11/301,094
- ; CURRENT FILING DATE: 2005-12-12
- ; PRIOR APPLICATION NUMBER: 60/636,390
- ; PRIOR FILING DATE: 2004-12-15
- ; NUMBER OF SEQ ID NOS: 24
- ; SOFTWARE: PatentIn version 3.3
- ; SEQ ID NO 6
- ; LENGTH: 1332
- ; TYPE: PRT
- ; ORGANISM: Homo sapiens
- US-11-301-094-6

Query Match 100.0%; Score 6909; DB 6; Length 1332;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVCLSN	60
QY	61	TPARKITGDHLLIQKICPRLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLLIQKICPRLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF	120
QY	121	VNLHCHTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDSCSRVRVPAA	180
Db	121	VNLHCHTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDSCSRVRVPAA	180
QY	181	ATLAVGTCMGVYGSALCNAQRWLNFGDGTGNGLAPLDDITFHLLEPCQAVGSGIOPNEG	240
Db	181	ATLAVGTCMGVYGSALCNAQRWLNFGDGTGNGLAPLDDITFHLLEPCQAVGSGIOPNEG	240
QY	241	ARCNSQGGDDVATCSQDCAASCPAIAFPQALDFTFYLGQMPGSLVLIILGCSFVAVVTI	300
Db	241	ARCNSQGGDDVATCSQDCAASCPAIAFPQALDFTFYLGQMPGSLVLIILGCSFVAVVTI	300
QY	301	LLVGRFVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV	360
Db	301	LLVGRFVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV	360

361	IPVVALAAGLVFTELTTDPVELWSAPNQARSEKAFHDQHFPGPFRTNQVILTAPNRSSY	420	APPLICANT: Dong, Jianli
361	IPVVALAAGLVFTELTTDPVELWSAPNQARSEKAFHDQHFPGPFRTNQVILTAPNRSSY	420	TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
361	IPVVALAAGLVFTELTTDPVELWSAPNQARSEKAFHDQHFPGPFRTNQVILTAPNRSSY	420	FILE REFERENCE: 02420/1201581-US1
421	RYDSLILGPKNFGSGLDLDDLLELELOERLHLQVWSPQARQNI SLQDICYAPLNPNT	480	CURRENT APPLICATION NUMBER: US/11/270,796
421	RYDSLILGPKNFGSGLDLDDLLELELOERLHLQVWSPQARQNI SLQDICYAPLNPNT	480	CURRENT FILING DATE: 2005-11-08
481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTTALAL	540	PRIOR FILING DATE: 2004-07-30
481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTTALAL	540	NUMBER OF SEQ ID NOS: 33
541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLFEM	600	SOFTWARE: PatentIn version 3.3
541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLFEM	600	SEQ ID NO 22
601	RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV	660	LENGTH: 1359
601	RAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV	660	TYPE: PRT
661	MYDSKATILGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE	720	ORGANISM: Homo sapiens
661	MYDSKATILGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE	720	US-11-270-796-22
721	YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLA	780	Query Match 99.5%; Score 6872.5; DB 6; Length 1359;
721	YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLA	780	Best Local Similarity 97.9%; Pred. No. 0;
781	ILDFLLQWSAFVALLSLSKQBEASRLDVCCCKVQOELPPPGQEGGLLGFQKAYAPFL	840	Mismatches 2; Indels 27; Gaps 1;
781	ILDFLLQWSAFVALLSLSKQBEASRLDVCCCKVQOELPPPGQEGGLLGFQKAYAPFL	840	Matches 1330; Conservative 0;
841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDVFLFNRYFVGAPV	900	
841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDVFLFNRYFVGAPV	900	
901	YFVTTLVGNFSSEAGMAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWVDVDFIDW	960	
901	YFVTTLVGNFSSEAGMAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWVDVDFIDW	960	
961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCWSITMGSVRPSVEQFHXYLPWFNDPR	1020	
961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCWSITMGSVRPSVEQFHXYLPWFNDPR	1020	
1021	NIKCPKGLAAYSIVNLTSQVTLASRFWAYHKPLKNSODYTEALRAARELAANITADL	1080	
1021	NIKCPKGLAAYSIVNLTSQVTLASRFWAYHKPLKNSODYTEALRAARELAANITADL	1080	
1081	RKVPCTDPAFVFPVTTINVEYQYLITLPEGLPMLSLCLVPTFAVSCLLGLDLRSGLL	1140	
1081	RKVPCTDPAFVFPVTTINVEYQYLITLPEGLPMLSLCLVPTFAVSCLLGLDLRSGLL	1140	
1141	NLLSTVMILVDTGFMALWDISYNVSLNINLVSAGMSVFVSHITRSPAISTKPTWLER	1200	
1141	NLLSTVMILVDTGFMALWDISYNVSLNINLVSAGMSVFVSHITRSPAISTKPTWLER	1200	
1201	AKEATISMGSAVPAAGVMTNLPGLILVLGLAKAQLIQIIFPFLNLLITLLGLHLGVLPLV	1260	
1201	AKEATISMGSAVPAAGVMTNLPGLILVLGLAKAQLIQIIFPFLNLLITLLGLHLGVLPLV	1260	
1261	ILSYVGPVNPALAEQKRAEEAANAAMVASCNHPHSRVSTADNIYVNHSPFEGSIKAGA	1320	
1261	ILSYVGPVNPALAEQKRAEEAANAAMVASCNHPHSRVSTADNIYVNHSPFEGSIKAGA	1320	
1321	ISNFLPNNGRQF 1332		
1321	ISNFLPNNGRQF 1332		
RESULT 2			
US-11-270-796-22			
; Sequence 22, Application US/11270796			
; GENERAL INFORMATION:			

Db 781 ILDFLLQMSAFVALLSDSKRQASRLSDVCCVKKPQLPPGQEGLLGLFFOKAYAPFL 840  
Qy 841 LHMTRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGPV 900  
Db 841 LHMTRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGPV 900  
Qy 901 YFVTTLYGNFSSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLEIPAASSWDDPIDW 960  
Db 901 YFVTTLYGNFSSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLEIPAASSWDDPIDW 960  
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGSVRPSVEQFHLYLPWFNDRP 1020  
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGSVRPSVEQFHLYLPWFNDRP 1020  
Qy 1021 NIKCPKGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053  
Db 1021 NIKCPKGLAAYSTSVNLTSDGQVLDVAILSPRLEYSGTISAHCNLYLLDSASRFMAYH 1080  
Qy 1054 KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1113  
Db 1081 KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1140  
Qy 1114 FMLSCLVPTFAVSCULLGLDLSGLNLLSIYMLVDTVGFMAWLDISYNAVSLINLVS 1173  
Db 1141 FMLSCLVPTFAVSCULLGLDLSGLNLLSIYMLVDTVGFMAWLDISYNAVSLINLVS 1200  
Qy 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGLVLGLAKAQ 1233  
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGLVLGLAKAQ 1260  
Qy 1234 LIQIPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPAALALQKRAEBAVAVMVASCP 1293  
Db 1261 LIQIPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPAALALQKRAEBAVAVMVASCP 1320  
Qy 1294 NHPSRVSTADNIYVNHSPSGIKGAGAINFLPNNGROP 1332  
Db 1321 NHPSRVSTADNIYVNHSPSGIKGAGAINFLPNNGROP 1359

RESULT 3  
US-11-270-796-3  
; Sequence 3, Application US/11270796  
; GENERAL INFORMATION:  
; APPLICANT: Dong, Jianli  
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND  
; TITLE OF INVENTION: RESTORATION OR MIMICRY PF p16 Ink4a ACTIVITY  
; FILE REFERENCE: 02420/1201581-US1  
; CURRENT APPLICATION NUMBER: US/11/270, 796  
; PRIOR FILING DATE: 2005-11-08  
; PRIOR APPLICATION NUMBER: 60/592,592  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 3  
; LENGTH: 1333  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-11-270-796-3

Query Match 78.3%; Score 5407; DB 6; Length 1333;  
Best Local Similarity 77.2%; Pred. No. 0;  
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy 4 AGLRGWLLALLRLAQSPYTIHQPGYCAFIDECKQPELSGSLMTLSNVSCLSNTPA 63  
Db 3 AAQWGLWALLNSAQSLYTPTHKAGCTFYECGKNPELSGSLTSLSNVSCLSNTPA 62  
Qy 64 RKTIGDHLTLLOKICPRLTYGN-TOACCSAKOLVSLEASLSITKALLTRCPACSDNFVN 122  
Db 63 RHVTGDHALLQRVCPRLTNGENDTYACCSTKQLVSLDSSLITKALLTRCPACSENFFVS 122  
Qy 123 LHCNTCPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAAT 182

Db 123 IHCNTCPNQSLFINVTRVQDPQQLPAVVAYEAFYQHSFAEKAYESCSRVRIPAAAS 182  
Qy 183 LAVGTWCGVYGSALCNAQRLNFGQDTGNGLAPLDITTEHLLEPGQVGSQIQLNEGVAR 242  
Db 183 LAVGSMCGVYGSALCNAQRLNFGQDTGNGLAPLDITTEHLLEPGQALADMKPLDGGKITP 242  
Qy 243 CNEQSGDDVATCSCODCAASCPAIAARPOALDSTFYLGQMPGSLVLIIILCSFAVVTILL 302  
Db 243 CNEQSGEDSAACSCQDCAASCPVIPPALRPSFYMGWPGWMLALIIIPFAVFLLSVVL 302  
Qy 303 VGFVAPADPKSNWDPKGTSLSDKLSFSTHTLLQFQCGWGTWASWPLTILIVSVIP 362  
Db 303 VYLRVASNRNKNKTAGSQEAPNLPKRRFSPHTVLGRFFESWGTVRASWPLTVLALSFIIV 362  
Qy 363 VVALAAGLVFTELTTPDVELWSAPNSQARSEAKAFHDOHFGPPFPRTNQVLTAPNRSYRY 422  
Db 363 VIALSVGLTFIELTTPDVELWSAPKQARKEAFHDEHFGPPFPRTNQIIFVAKNRSYKY 422  
Qy 423 DSIILGPKNFSGILDLDLLELLELQERLRLHQLQVMSPEAQRNISLQDICVAPLNPNNTSL 482  
Db 423 DSIILGPKNFSGILSLDLLELLELQERLRLHQLQVMSPEAQRNISLQDICVAPLNPNNTSL 482  
Qy 483 YDCCINSLLOYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALALSC 542  
Db 483 TDCCVNSLLQYFQNNHTLLLTANQTLNGQTSVDMKDHFLYCANAPLTFKGTALALSC 542  
Qy 543 MADYGAPVFPFLAIGYKGDYSEAEALJMTESLNNYPAGDPRLAQAKWEAEFLBEMRA 602  
Db 543 IADYGAPVFPFLAIGYKGDYSEAEALJMTESLNNYPAGDPRLAQAKWEAEFLBEMRA 602  
Qy 603 FORMMAGFQVFTFAERSLEDEINRTTABDLPFATSYIVFIYISIALGSSYSWSRVMV 662  
Db 603 FQSTADKQIAFSAERSLEDEINRTTIQDLFPFALSYLIVFIYISIALGSSYSWSRVAV 662  
Qy 663 DSKATLGLGGVAVVLGAVMAAGFPFSGYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQ 722  
Db 663 DSKATLGLGGVAVVLGAVMAAGFPFSGYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQ 722  
Qy 723 RLPRPGEPREVHIGRALGRVAPSMLLCSLSAICFFLGAITPMPAVRFTALTSLGAVIL 782  
Db 723 RLPRPGEOREAHIGRTLGSVAPSMLLCSLSAICFFLGAITPMPAVRFTALTSLGAVIL 782  
Qy 783 DELLQMSAFVALLSDSKRQASRLSDVCCVKKPQLPPGQEGLLGLFFOKAYAPFLH 842  
Db 783 DELLQMTAFVALLSDSKRQASRLSDVCCVKKPQLPPGQEGLLGLFFOKAYAPFLH 842  
Qy 843 WITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYF 902  
Db 843 RPIRPVLLFLVLFGANLYLNCNITSVGLDQDLALPKDSYLLDYFLNRYLEVGPVYF 902  
Qy 903 VTTLYGNFSSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLEIPAASSWDDPIDWT 962  
Db 903 DTTSGYNFSTEAGMNAICSSAGCANNFSTQKIQYATEPPEQSYLEIPAASSWDDPIDWT 962  
Qy 963 P-SSCCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGSVRPSVEQFHLYLPWFNDRP 1021  
Db 963 PSSSCCRLYITRGPHKDEFCPSTDTSFNCLNCKMNTLGPVVRTTQFHLYLPWFNDRTN 1022  
Qy 1022 IKCPKGLAAYSTSVNLTSDGQVLAASRFMAYHKLKNSQDYTEALRAARELANITADLR 1081  
Db 1023 IRCPKGGLAAVYTSVNLSSDQIIASQFMAYHKPLRNSQDPTTEALRASRLAANITAEUR 1082  
Qy 1082 KVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCULLGLDLSGLN 1141  
Db 1083 KVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCULLGLDLSGLN 1142  
Qy 1142 LLSIWMILVDTVGFMAWLDISYNAVSLINLVSACMSVEFVSHITRSPAISTKPTWLER 1201  
Db 1143 LLSIWMILVDTVGFMAWLDISYNAVSLINLVSACMSVEFVSHITRSPAISTKPTWLER 1202  
Qy 1202 KEATISMGSAVPAVAMTNLPGLVLGLAKAQLIQIPFRNLNLTLLGLLHGLVFLPV 1261

Db 1203 KDATIFMGSAVAGVAMTNPFGIILITLGPQAQLIPIFFRLNLLITLGLHLGLVFLPV 1262  
Qy 1262 LSYVGPVNPALAEQKAEAAVAVMVASCPNHPSTADNLYVNHSPFGS-1KGAGA 1320  
Db 1263 LSYLGPVNPALVLEEKATEA-AMVSEPCQPPFPADANTSIDVYNGFNPEPIELNA 1321  
Qy 1321 ISNPLPNNGRQF 1332  
Db 1322 ASSSLPKSDQKF 1333

RESULT 4  
PCT-US06-05584-822  
; Sequence 822, Application PC/TUS0605584  
; GENERAL INFORMATION:  
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION,  
; TITLE OF INVENTION: ASSESSMENT, PREVENTION, AND THERAPY OF CANCER  
; FILE REFERENCE: DFS-064.25(25992-6425)  
; CURRENT APPLICATION NUMBER: PCT/US06/05584  
; CURRENT FILING DATE: 2006-03-02  
; PRIOR APPLICATION NUMBER: 60/590,064  
; PRIOR FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: 60/654,227  
; PRIOR FILING DATE: 2005-02-17  
; NUMBER OF SEQ ID NOS: 848  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 822  
; LENGTH: 1278  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US06-05584-822

Query Match 34.8%; Score 2402.5; DB 1; Length 1278;  
Best Local Similarity 39.8%; Pred. No. 1.5e-185; Indels 103; Gaps 25;  
Matches 522; Conservative 237; Mismatches 450;

Qy 7 RGLWALLRLRLAQSEPTTTHQPGYCAFYDECG-----KNPELGSGLMTLSNVCSLS 59  
Db 4 RGLALGLLLLLCPAQVFSG-----SCVWYEGGIAYGDKRYNCEYSG----- 46

Qy 60 NTPARKITGDHLILLOKICPLRYTGNTQACSAKQVLSLEASLTITKALLTRCPACSDN 119  
Db 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKDNLQPLQLSRCPSCFYN 103

Qy 120 FVNLHCHNTCSNQSLEFINVTR---VAOLGAGQLPAVVAEAFYQHSFAEQSDYSCSRV 175  
Db 104 LMLFCELTCSRQSQFLNVATEDYDPTNQTKNVVELQYVYVQSQFANAMYNACRDV 163

Qy 176 RVPAAATLAVGTGCMGVYGSALCNAQRWLNFGQDTGNGLAPLDIT-----FHLLEPGQA 228  
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIETMFKNDNGQAPFTITPVFSDFPVH----- 215

Qy 229 VSGIQPLNEGVARCNESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279  
Db 216 ---GMBPMNATKGCDESDEVETAPCSQDCSVCGPKQPPPPPPAPWTILGLDAMVIM 272

Qy 280 QMPSGLVLIILCSVFAVITILL-----VGRFVAPADKSKMVDPPKGTSL 326  
Db 273 WITMAFLVIFGFAFVAVCYRKYRYFSEVTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHLLGQFQGWGTWASWFLTILVLPVVALAAGLVTELTDTDELWLSAP 386  
Db 325 DPVSAAPGECRLRLFRWGSFCVRNPGCVIFFSLVFITACSSGLVFRVTRVTPNVDLWSAP 384

Qy 387 NSQARSEKAPHDQHFQPFRTNQVILTAPNRESSVRYDLSLLGPK-NFSGILDLDLLELL 445  
Db 385 SSQARLEKEYDQHFQPFRTNQVILTAPNRESSVRYDLSLLGPK-NFSGILDLDLLELL 444

Qy 446 ELQERLRLHQLWSPQEAQRNLSQDICVAPLNPNTSLYDCINSLSLQYFQNNRLLTLLTA 505  
Db 445 DLQIAIEN--ITASYNDETTLQDICLAPLSPYNT---NCTILSVNLNFGNSHSLVDHKK 499

Qy 506 NOTIMGOTSQVDKDHFLYCANAPLTFKDTALALSCMADYAGAPVPPFLAIGCYKGDYS 565  
Db 500 GDDFF---VYADYTHFLCYCRAPASLNDTSLDCLGTFGGFVFPWLVGLGGDDQNYN 556

Qy 566 EAEALIMTFSLNNYPAGDRLAOKLWEEAFLEEMRAFORRMAGMFOVTTAERSLDEI 625  
Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFKYKN---PNLTISFTAERSIEDEL 613

Qy 626 NRTTAEDLPFAYSYIVIFYLSIALGSSYSRVMVDKATLGLGVAVVLGVMAAMG 685  
Db 614 NRESDSVETWISYAIMFLYISLALGHKICRRLVDSKVSIGIAGILVLSVACSLG 673

Qy 686 FFSYLGRSLVILQVFPFLVLSVGADNIIIFVLEYQLRPRRPEPEVHIGRALGVAP 745  
Db 674 VFSYIGLPLTLVIEVIFPLVAVGVNIIIFILVOAYORDERLOETDUQQLGRVGEVAP 733

Qy 746 SMLLCSLSEACFFLPGALTMPAVRTPALTSGLAVIDLFLQMSAFVALLSLDSKROEAS 805  
Db 734 SMFLSSSEVAPFLGALSVPVAVHTSLFAGLAVFIDFLQITCFVSLGLLDIKROEKN 793

Qy 806 RLDVCCVCKQELPPQO-GEGLLLGPFQKAYAPFLHWHITRQVVLVLLFLALFGVLSYM 864  
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKLDWMRPVIAIFVGVLSFSIAVL 853

Qy 865 CHISVGLDDELALPKDSYLDYFLFLNRYFEVGAAPVFTVTLGVNFSSEAGMNAICSSAG 924  
Db 854 NKVDIGLDSLSMPDDSYMVDYFKSISOYLHAGPPVTVLEEGHDTISSKQNNVCGMG 913

Qy 925 CNPFSFTQKIOYATEFPEQSYLAIPASSWVDDDFIDWLTP--SSCRLYISGPNKDKFCPST 983  
Db 914 CNNDLSVQOIFNNAQLDNYTRIGFAPSWDIDYFDWVKPQSSCCRV---DNITDOFCNAS 970

Qy 984 VNSLNLKNCMSIT-MGSVRPSVEQFHLYLPMFLNDRPNKICPKGGLAAYSYVN--LTS 1040  
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPCRGKGHAAYSSAVNILLGH 1029

Qy 1041 DGQVLASRPMAYHKLKNSQDYTEALRAARELANITADLRKVCTDPAPFVFPVTITNV 1100  
Db 1030 GTRVGATYFMTYHTVLTQSDAFIDALKKARLIASNT-ETWINGS--AYRVFYSVFPV 1086

Qy 1101 FVEQVLTILPGLFMLSCLVPTFAVSCLLGLDLRSLGLNLLSIVMLVDTVGFMAIWD 1160  
Db 1087 FVEQVLTIIDTIFNLGVSGLGALFLVTMVLVGLCELWSAVIMCATIYMLVNMFGVWLWG 1146

Qy 1161 ISYNASVLINLVASVMSVEFVSHITSFALSTPTWLERAKEATISMGSAVEAGVAMTN 1220  
Db 1147 ISLNAVSLVNLVMSGISEVFCSHITRAFTVSMKGSRYERAEALAHMGSSVFSGITLTK 1206

Qy 1221 LPGAIVLGLAKAQLIQIIPFFRLNLLITLGLHLGLVFLPVLISYVGPVNDPA 1272  
Db 1207 FGGIVVLFAKSAQIQIIFFRNYLAWLGLGTHGLIFLPLVLLSYIGPSVKA 1258

RESULT 5  
US-11-191-274A-129  
; Sequence 129, Application US/11191274A  
; GENERAL INFORMATION:  
; APPLICANT: Applera Corporation  
; APPLICANT: Bruno DOMO  
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF  
; FILE REFERENCE: CL001536PROV  
; CURRENT APPLICATION NUMBER: US/11/191,274A  
; CURRENT FILING DATE: 2005-07-28  
; NUMBER OF SEQ ID NOS: 334  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 129  
; LENGTH: 1278  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-191-274A-129

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;  
Best Local Similarity 39.8%; Pred. No. 1.5e-185;



```

445 DLQAIEN--ITASYDNETHLQDCLAPLSPYNT---NCTILSVLNYPQNSHVLDDHKK 499
506 NQTLMGQTSQVMDKWHFLYCANAPLTFKGTALALSCMADYGAAPVFPFLAIGYKGYDS 565
500 GDDEF--VYADYHFLYCVRAPASLNDTSLHDPCLGTFGGVFPVFLVGGYDDQNYN 556
566 EAEALIMTFLSNYPAGDPLAQAALWEAEFLERAFORMAGMFQVTFPFAERSLEDEI 625
557 NATALVITFPVNNYNDTEKLOQAQAEKEFINFVKYKN---PNLTISFTAESIEDEL 613
626 NRTTAEPLIPATSYIVIFLYISLALGYSYSSRWVMDVKATLGLGVAVLGVAVMAAMG 685
614 NRESDDVFTVVISYAMFLYISLALGHKICRLLVDSKVSIGIAGILVLSVACSLG 673
686 FFSYLGIRSLVTLQVVPFLVSVGADNIFIFVLEYORLPRRPGEPREVIHGRALGRVAP 745
674 VFSYIGLPLTLIVIEVIFLAVGVDNIFILVQAYQDERLQGETLDQQLGRVLGEVAP 733
746 SMLLCSLSRAICFELGALTPMPAVRTFALTSLGLAVILDELQMSAFVALLSLSKROEAS 805
734 SMLSSFSETVAFFLGLSVNPAVHTFSLFAGLAVFIDELQITCFVSLGLDITKROEKN 793
806 RLDVCCVCKPQELPPGQ--GEGLLGFFQKAYAPFLLHWITRGVLLFLALFGVSLYSM 864
794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMRPIVIAIFGVLSIAVL 853
865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAAPVFTVTLGYNPFSSAGNNAICSSAG 924
854 NKVDIGLDQSLMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKQNMVCGMG 913
925 CNNFSFTOKIYATEPPEQSYLAIPASSWVDDFDLWLT--SSCCRLYISGNPKDFCPS 983
914 CNNDLSVQQLFNAQAQDNITRGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
984 VNSLNCCLKNCMSIT--MGSVRPSVEQPHKYLFWFLNDRPNIKCPKGLAAAYSTSVN--LTS 1040
971 VVDPACVR--CRLTPGKQRPQGGDFMRFLPWFSLDNPNPKCGKGHAAYSAVAILLGH 1029
1041 DGOVLASRWYHKKPLKNSQDYTEALRAARELANITADLRKVPGTDPFAFEPYITNV 1100
1030 GTRVGATYFMTVHTVLTQTSADFIDALKKARLIASNT--ETMGINGS--AYRFPYSVFV 1086
1101 FFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLSIWMILVDTVGMALWD 1160
1087 FFEQYLTIDDDIFNLGVSIGLAFVTVMLGCELWSAVIMCATIANVLNNFVGMWLMG 1146
1161 ISYNAVSLNLVSAGVMSYEFVSHITRSPAITKPTWLERAKEATISMSGSAFAGVAMTN 1220
1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVSMKGSVRERASEALAHMSSVFSGITLTK 1206
1221 LBGILVLGAKAQLOIIFFRNLNLTITLILGLHGLVFLPVILSYVGPVWPA 1272
1207 FGGIVLAFAKSQIIFOIFPFRMYLAMVLLGATHGLIFLPLVLLSYIGPSYNKA 1258

```

RESULT 7

```

US-11-385-692-2400
; Sequence 2400, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385, 692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2400
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2400

```

```

Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMILLWALLLLAOSPEVYTHIQPCYCAFYDECG-----KNPELSGLTSLTNSVCSLS 59
DB 4 RGLALGALLLLLCAPQVFSQ-----SCVWYEGCGLAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDLHLILLOKICPRLYTPGNTQACCSAKQLVSLSEASLITKALLTRCPACSDN 119
DB 47 --PKPLPKDGYDLVQELCPGFFFG--NVSLCDDVQLQQLTKDNLQQLPLQFLSRCPSCFVN 103
QY 120 FVNLHCNTCSPNOSLFINVTR---VAQLCAGQLPAVYAEAFQHSFAQSVDSCSRV 175
DB 104 LNLUFCELTCSPROSQFLNVTATEDYVDPVNTQNTKTNVKELOQYVVGOSFANAMYNACRDV 163
QY 176 RVPAATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDIT-----PHLLEPQA 228
DB 164 EAPSSNDKALGLLCGKOADA--CNATNMEYMFNKDNGQAPFTITPFSDFPVH----- 215
QY 229 VSGIQPLNEGVARNCSQGGDDVATCQDCQCAASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GMPEMNNATKGCDSVDEVTAPECSQDCSIVCGPKPQPPPPAPWTLGLDAMYIM 272
QY 280 QMPGSLVLIILCISVEAVVTILL-----VGFRVAPARDKSKWVDPKKGTSL 326
DB 273 WITWAFLLVFPFAFFAVMCYRKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
QY 327 DKLFSFTHLLGQPFQCGMTWVASWPLTILVLSVIPVVALAAGLVFTLETTDPVELWSAP 386
DB 325 DPVSAAPFEGCLRLLFTRWGSFCVENPGCVIFFSLVFITACSSGLVFRVVTNPVDLWSAP 384
QY 387 NSQARSEKAFHQHGFPPFTNTQNVILFAPNKRSSVRYDSLLGPK--NFSGILDOLLLELL 445
DB 385 SSOARLEKEYDQHFHGFPRTEQLIIRAPLTKHIYQIPYPSGADVPPGPDIDQILHQVL 444
QY 446 ELQERLRLHQLVMSPEAQRTISLODICVAPLNPDNTSLDYDCINSLLOYFQNNRTLLLT 505
DB 445 DLQAIEN--ITASYDNETHLQDCLAPLSPYNT---NCTILSVLNYPQNSHVLDDHKK 499
QY 506 NQTLMGQTSQVMDKWHFLYCANAPLTFKGTALALSCMADYGAAPVFPFLAIGYKGYDS 565
DB 500 GDDEF--VYADYHFLYCVRAPASLNDTSLHDPCLGTFGGVFPVFLVGGYDDQNYN 556
QY 566 EAEALIMTFLSNYPAGDPLAQAALWEAEFLERAFORMAGMFQVTFPFAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKLOQAQAEKEFINFVKYKN---PNLTISFTAESIEDEL 613
QY 626 NRTTAEPLIPATSYIVIFLYISLALGYSYSSRWVMDVKATLGLGVAVLGVAVMAAMG 685
DB 614 NRESDDVFTVVISYAMFLYISLALGHKICRLLVDSKVSIGIAGILVLSVACSLG 673
QY 686 FFSYLGIRSLVTLQVVPFLVSVGADNIFIFVLEYORLPRRPGEPREVIHGRALGRVAP 745
DB 674 VFSYIGLPLTLIVIEVIFLAVGVDNIFILVQAYQDERLQGETLDQQLGRVLGEVAP 733
QY 746 SMLLCSLSRAICFELGALTPMPAVRTFALTSLGLAVILDELQMSAFVALLSLSKROEAS 805
DB 734 SMLSSFSETVAFFLGLSVNPAVHTFSLFAGLAVFIDELQITCFVSLGLDITKROEKN 793
QY 806 RLDVCCVCKPQELPPGQ--GEGLLGFFQKAYAPFLLHWITRGVLLFLALFGVSLYSM 864
DB 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMRPIVIAIFGVLSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAAPVFTVTLGYNPFSSAGNNAICSSAG 924
DB 854 NKVDIGLDQSLMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKQNMVCGMG 913
QY 925 CNNFSFTOKIYATEPPEQSYLAIPASSWVDDFDLWLT--SSCCRLYISGNPKDFCPS 983
DB 914 CNNDLSVQQLFNAQAQDNITRGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCCLKNCMSIT--MGSVRPSVEQPHKYLFWFLNDRPNIKCPKGLAAAYSTSVN--LTS 1040

```

```
Db 971 VDPACVR-CRPLTPEGKQRQGGDFMRFLPMFLUSDNPNPKCGRGHAAAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNV 1100
Db 1030 GTRVGATVMTYHTVLTQTSADFIDALKKARLIASNVT-ETMGINGS--AYRVFPYSVFYV 1086
Qy 1101 FYEQYLILBEGFLPMLSLCLVPTTFAVSCLLGLDLRSGLLNLLSIVMLVDTVTFMALWD 1160
Db 1087 FYEQYLTIIDDTIFNLGVSGLGAIFLVMTVLGCELWSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGVMSVEFVSHITRSEALSTKPTWLERAKEATISWGSNAVFAGVAMTN 1220
Db 1147 ISLNAVSLVNWSCGLSVFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIIFFRNLNLLITLGLLHGLVFLPVLTSYVGPDPVNP 1272
Db 1207 FGGIVVLAFAKSQIQIFVFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVKA 1258

RESULT 8
US-11-385-692-2401
; Sequence 2401, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385.692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2401
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2401
```

```
Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGLWLLALLRLAQSFPYTHIOPGCAFYDECG-----KPELSGSLMTLSNVCSLS 59
Db 4 RGLALGLLLLLCPAQVFSG-----SCVMYGEGLAYGDKRYNCEYSG----- 46

Qy 60 NTPARKITGDHLILLOKICPELYTGNTQACCSAKOLVSLASLSITKALLTFCPCSDN 119
Db 47 --PPKPLPKDGYDLVDELCPGFFG--NVLCCDVRLQTLKONLQPLQLGRCPSCFYN 103

Qy 120 FVNLHCHNTCSNPQSLFINVTR----VAQLGAGQLPAVVAYBAFYQHSFAEQSYDSCSRV 175
Db 104 LNLNLFCELTCSPQSQFLNVTATEDYVDPVTNQTKNVXELQYVYVQCSFANAMYNACRDV 163

Qy 176 RYPAATLAVGTWCGVYGSGALCNAQRLWNPQDGTGNGLAPLIDT-----PHLLPQGA 228
Db 164 EAPSSNDKALGLLIGCDADA-CNATWIEYMFNKONGQAPFTITPVFSDFPVH----- 215

Qy 229 VSGGIPLNEGVARCNESQDDVATCSCODCAASCPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGDSEVDEVTAPCQDCQDSIVCGPKPQPPPPPPAPWTILGLDAMYIM 272

Qy 280 QMPGSLVLIIILCSVPVATILL-----VGRFVAPARDKSKWDPKKGTSLS 326
Db 273 WITYMAFLIVFFGAFAVVCYRKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFGGNGTGWASWPLTILVSVIPVVALAAGLVFTLTDDPVELWSAP 386
Db 325 DPVSAFAEGCLRLRLFRWGSFCVRNPGCVIFFSLVFTACSSGLVFRVTTNPDVLSAP 384

Qy 387 NSQARSEKAFHQHFGPFRTHNQVILTAPNRSRYDLSLLGPK-NFSGILDLLELL 445
Db 385 SSQARLEKEYFQDHFGPFRTHNQVILTAPNRSRYDLSLLGPK-NFSGILDLLELL 444
```

```
Qy 446 ELQERLRHLQVMSPEAQRNLSLDQICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASYNDETTLQDICIPLAPSPYNT---NCTILSVNLVFNQSHSVLHKK 499
Qy 506 NQTLMGQTSQVDKDHFLHCANAPLTFKGTALALSCMADYGAPVFPFLAIGGYKGDYS 565
Db 500 GDDFF--VYADYHTFLCYCRAPASLNDTSLLDHPCLTGFGGPFVFWLVLGGYDDQYN 556
Qy 566 EEAALIMTFSINNYPADGPRLAQAKLWEAEFLREMAFORRMAGMFOVTTAERSLDEI 625
Db 557 NATALVITFPVNNYNDTEKLRQAQWEKEFINFKYKN---PNLTISTATERSIEDEL 613
Qy 626 NRTTABDLPIFATSYIVIFYISLALGSYSSWSRVMVDKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDDVFTVVISYAIMPLYISLALGHIKSCRRLLVDSKVSIGIAGILLVSSVACSLG 673
Qy 686 FFSYLGIRESLVLQVVPFLVLSVGDADNIFIVLEVRQRPRLRRGEPREVIHIGALRVAP 745
Db 674 VFSYIGLPLTLVIEVIFPLVLAAGVDNIFILVQAYQORDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSAICFFLGCALTMPMPAVRTFALTSLGLAVILDFLLQMSAFVALLLSLDSKREAS 805
Db 734 SMFLSFSFTVAFFLGALSVMPPAVHTPSLFAGLAVFIDFLQITCFVSLGLDIIKKQEN 793
Qy 806 RLJVVCCVRPQELPPPGQ--GEGLLLGFFQKAYAPFLHMTTRGVWLLLFALFGLSVLSYM 864
Db 794 RLIDIFCCVRGAEDGTSVQASESCLPFRFFKNSYSPLLKDWMRPIVITAFVGLVSFSI AVL 853
Qy 865 CHISVGLDQELALPKDSYLLDLYFLNRYFEVGAPVYVTTGLYGNFSSBAGMAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYNDYFKSISQYLHAGPPVYFVLEEGHDYTSKQGNMVCGMG 913
Qy 925 CKNFSPKIOVATEPEEQSYLAIPASSWVDDFDLWLT-SSCCRLYISGPNKDKCPCST 983
Db 914 CNDLSLVQQLFNAQLDNTTRIGFAPSSWIDYDFDMVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNCNLCNSIT-MGSVRPSVEQFHKYLPFLNDRPNIKCPKGLAAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRQGGDFMRFLPMFLSDNPNPKCGRGHAAAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNV 1100
Db 1030 GTRVGATVMTYHTVLTQTSADFIDALKKARLIASNVT-ETMGINGS--AYRVFPYSVFYV 1086
Qy 1101 FYEQYLILBEGFLPMLSLCLVPTTFAVSCLLGLDLRSGLLNLLSIVMLVDTVTFMALWD 1160
Db 1087 FYEQYLTIIDDTIFNLGVSGLGAIFLVMTVLGCELWSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGVMSVEFVSHITRSEALSTKPTWLERAKEATISWGSNAVFAGVAMTN 1220
Db 1147 ISLNAVSLVNWSCGLSVFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIIFFRNLNLLITLGLLHGLVFLPVLTSYVGPDPVNP 1272
Db 1207 FGGIVVLAFAKSQIQIFVFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVKA 1258

RESULT 9
US-11-301-094-2
; Sequence 2, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Smith, Marsha
; APPLICANT: Levitan, Diane J
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Ratent in version 3.3
; SEQ ID NO 2
; LENGTH: 1383
```

```
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-301-094-2

Query Match      19.4%; Score 1341.5; DB 6; Length 1383;
Best Local Similarity 27.5%; Pred. No. 2e-99;
Matches 369; Conservative 257; Mismatches 561; Indels 155; Gaps 40;

Qy 76 KICPRLVYGPNTQACCSAKQLVLSLEASITKALLTRCPACSDNFVNLHCHTCSNQSL 135
Db 63 EFCPHLLTGDN-KLCCTPSAEGLTKQIAQARHLILGRCPSCFDNPAKLWCBFTCSNQD 121
Qy 136 FINTRVAQL--GAGQLPAVAYEAF-----YQHS--FAEQSYDSCSRVRPAAATLAVG 186
Db 122 FVSIEMKPIEBKGGFTPEYQPAEAYVNTVEIRLSTDAEGMFSCKDVTFGGQPALRV- 180
Qy 187 TMCGVYGSALCAQORWLNFOGDTGNGL-APLIDITFHLLPEQAVGSGIQP-LNBSGVARCN 244
Db 181 -MC---TSTPTCLTNWLEFICTQLNDLNIPIHTKFLLYDPIKTPPSDRSTVMNVNFTGCD 236
Qy 245 ESQGDVATCSQDCAASCRAIPARQALDSTFYLGOM-----PGSLVLIILLCSFVAVVT 299
Db 237 KSARVGWPACTSEC--NKEBYANLIDLDGKTSQTCNVHGIACLNIFVMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARDKSMVDPKKGTSLDKLSFSTHTLLGQFF 341
Db 295 LLCVGFVFTSDEDTNLRQTSQGEESPRNRK-----RTGAWI-----HNFMENNA 342
Qy 342 QCMGTWASWPLTILVLSVIPVALAAGLVFTELTDPVELWSAPNSQARSEKAPHQHF 401
Db 343 RDIGNMAGNPKSHFFICGAVLIFCLPGMIYHKESTNVVDMWSSPRSRARQEEVFNANF 402
Qy 402 GPFFRTNQVILTAPNRSYRVDLSLLGPKNPSGILDLLELLELRLHQLRWLSPEA 461
Db 403 GRPQRYQOIMLL--SHRDFQSGKLYGP-----VFHKDIFBEELFIDILNAIKNISQSDG 455
Qy 462 QRNISLQICYAPLNPNTSLYDCCINSILQYFQNNRTLLLTANOT-----L 509
Db 456 -RTITLDDVCVRPMFG-----YDCLIMSPNTYFQNGKHLDMKSKETVSEDDDAFYF 510
Qy 510 MGQTSQVDKWHFLYCANAPLTFKDGATLALSCMADYGAIPVFPFLAIGYRGKOYSEAE 569
Db 511 SSEATTDEWNNHMAACIDQPMHSQK--TKSGLSCMGTYGSPAPNM-VFGKSTNHQAANS 567
Qy 570 LIMTSLNYPAGDPRLLQAQWEEAFLEEMRAFORRMAGMPQVTFTAERSLEIDNRTT 629
Db 568 IMWTLIVTQ--RTEPEIKQAELEWEKFLKFCKEYREKSPKVI-FSFMABERSITDEIEN 624
Qy 630 AEDLPFATSYIVIFLYISLALGSY-----SSMSRVMDSKATLGLGVAVVLGAVMAAM 684
Db 625 KDEIVTVVIALAFLGYVTFSLGRYFVCENQLWS--ILVHSRICLGMVSLVINLLSSFCSW 683
Qy 685 GFYSYLGRSSILVILQVVPFLVLSVGADNIFIVLEY--QRLP---RRPGEPREVHIGRA 739
Db 684 GIFSMFGIHPVKNALVQVFPVTVLLGVCHTFWVKYQAQRYVMPYMSPDQCPFI-VGMV 742
Qy 740 LGRVAPSMLLCSLSEACFFLGALTPMPAVRTALTSLGLAVILDFLOWSAFVALLSDS 799
Db 743 MAGTMPAMPSSSLGCAFSFFIGFTDLPAIRTFCLYAGLAVLIDVVLHCTIIFALFVMDT 802
Qy 800 KQAEASRLDCCCVKQPELPPQCGEGL-----LIG-----FFQKAVAPFLHWHI 844
Db 803 QRELNG-----KP-EFFFPVQIKDLLGAVLIGRQATDTFMQTFHFQVAPFLMHRM 853
Qy 845 TRGVWLLFLALFGVSLYSWCHI SVGLDQELALPKDSYLLDYFLFLNRYFVVGAPYFVT 904
Db 854 TRIITGIIIFASPIITVILSSKISVGFQDSMAFTKSYISTHRYLDKDFDVGPPVFPV 913
Qy 905 TLGYNFSBAGMNAICSSAGCNFNFSPTQKIQVATEPPEQSYYLAIPASSWVDDFDIWLTP 963
Db 914 DGELDWHRPDVQNKFTFPFGCSDTSFNGIMNAVGHTEQTYLSGENYMNWIDYLEWISK 973
Qy 964 SSSCRLYISGPNKDKECPSTVNSL-----NCLKNCM-----SITMGSV---RPSVEQPH 1009

974 SPCKVYVHDPN--TFCSTNRNKSALDDKACRTCMDFDYVANSYPKSSIMYHRPSIEVFY 1031
1010 KYLPWELNDPRNLIKCPKGGAAVSTVNLTSDQVLASRFMAYHKPL--KNSODYTEALR 1067
1032 RHURHFLEDTNSECVFGRASPKDAISFTSRGRIQASQPMTHKKLSISNSDFIKAMD 1091
1068 AARELAANTADLRKRVGTDPAFEVFPYTTITNVFYEQYLITILPEGLFMLSCLVPTPAVS 1127
1092 TARMVSRRLERSI-----DDTAHVPAYSKIPFFEQYSTIMPILITQLFITVVGVGII 1145
1128 CLLGLDLRSLNLNLSIVMILVDYTFGFMALWDISTNAVSLINLVSAGVMSVFEVSHITR 1187
1146 CVTLGIDVGAACAVICQVSNYFHVAFMYIFINPVNALSATNLVMSGGILIEFSVNVLK 1205
1188 SFAISKPTWLERAKEATISMGSAPGAVAMTNLPGILVLGLAKAQLIOIIPFFRLNLLIT 1247
1206 GYACSLRQRAKORAEESTVGSIGPIILSGPVVTVWAGSTMFLSGAHLQIITVYFKLFLITI 1265
1248 LLGLLHGLVLPVILSYVGPVNPALALEQKRAEEAVAAVMVASCPNHPHRSRVSTADNIYV 1307
1266 VSSAVHALIILPILAFGSGRSGHSGSSETSTNDNDQHDACVLS--PTAESHISNVERGIL 1323
1308 N-----HSFEGSINKGAGAI 1321
1324 NRPSLLDASHILDPLLKAEGGI 1345

RESULT 10
US-11-301-094-4
; Sequence 4, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JH06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-301-094-4

Query Match      15.1%; Score 1046; DB 6; Length 1274;
Best Local Similarity 25.7%; Pred. No. 1.6e-75;
Matches 345; Conservative 241; Mismatches 587; Indels 168; Gaps 46;

Qy 10 LLWALLLRLAQSEBPTTHQPGYCAFYDECCKNPFELSGSLMTLSNVSCLSNTPARKITGD 69
Db 14 VLFLLLIHLALCQ-----AKCMV-TECDGEEDSNHPPCKTNKSTYLPITVTRSLNPT 64
Qy 70 HLILLOKICPLRYLTGPN--TOACCSAKQLVLSLEASITKALLTRCPACSDNFVNLHCHNT 128
Db 65 YNARPEKYCSVLVQEEDEKAQVCCTELQKGMTDRISNAATILGSCPCFDNFAXLWCQFT 124
Qy 129 CSPNQSLFINTRVAQAGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAAATLAVGTM 188
Db 125 CSPDQSKEMKWET-----TGPKNVVVVKMEFKNRDFVEGLYESCRHTWFANGLAURLMSL 180
Qy 189 GQVYGSALCAQORWLNFOGDTG-NGLA---PLDITFHLLPEQAVGSGIQPLNEGVARN 244
Db 181 GQKVS-----FENFYGFMTGKWLQASIPINTEFQFSRMKNANWIPTP-----CH 225
Qy 245 ESQGDVATCSQDQD-----AASCPAIARPAQALDSTFYLGQMPGSLVLIILCSVFVAVTIL 301
Db 226 KSAGKVPACGNAIDCPTNAHQVLVDISKVEHLGTGVFHPFP--DFEWLLKICGLA-LTVL 283
Qy 302 LVGF-----RVAPARDKSMVDPKKGTSLDKLSFSTHTLLGQFFQG-----WGTWAS 350
```



Mon Apr 10 07:16:15 2006

APPLICANT: Ujwal, Manusha L.  
APPLICANT: Ma, Yunging  
APPLICANT: Chen, Rui-Hong  
APPLICANT: Ghosh, Malabika  
APPLICANT: Weng, Gezhi  
APPLICANT: Haley-Vicente, Dana  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
FILE REFERENCE: 823  
CURRENT APPLICATION NUMBER: US/10/461,673  
CURRENT FILING DATE: 2003-06-13  
PRIOR APPLICATION NUMBER: PCT/US02/29964  
PRIOR FILING DATE: 2002-09-19  
PRIOR APPLICATION NUMBER: US 10/245,014  
PRIOR FILING DATE: 2002-09-16  
PRIOR APPLICATION NUMBER: US 60/323,739  
PRIOR FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: PCT/US02/29636  
PRIOR FILING DATE: 2002-09-18  
PRIOR APPLICATION NUMBER: US 10/245,817  
PRIOR FILING DATE: 2002-09-16  
PRIOR APPLICATION NUMBER: US 60/323,349  
PRIOR FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: PCT/US02/29001  
PRIOR FILING DATE: 2002-08-09  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 17116  
SOFTWARE: pf\_FL\_genes Version 6.0  
SEQ ID NO 16739  
LENGTH: 967  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-461-673-16739

Query Match 8.5%; Score 588.5; DB 6; Length 967;  
Best Local Similarity 21.3%; Pred. No. 1.4e-38;  
Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;  
334 HT-----LLGQFFQGWGTWAGWPLTILVSVVVALAAGLVF--TELTDPVLSAP 386  
115 HTDLEGLSRTFQWLGWGAHPWTFLLAPLMLTAALGTGLYLPKDEEDLEEHTPV 174  
387 NSQARSEKAFHGHGPFRTNQVILTAENRSS-----YRYSLLGKPNFSGIL 436  
175 GSPAKAERRFVQGH-----FTTNDYSRFSASRRSTEANFVLSVSDSLDPAFVYS 230  
437 DLDLLELLELQERLRLHQLWSPEAQKNTSLQDIC--YAPLNPDTNSLYDCCI--NSLQ 492  
231 KLDGAVQDDURV-----AREKSQIQOVCARYAL-----CVPPNPILY 270  
493 YFQNNRTLLALTANQITLMGOTSQVDMKDFLYCANAPLTPKGTALALSCMADYGA-PVF 551  
271 AQVQNTLNL-----SSISFPAY-----NHRHPLY 296  
552 PFLAIGY-----KQDYSEAEALIMTFSNAPPAGDPRL-AQAKLWEEAFLEENRAF 603  
297 LTGFFGGLGSLGMLGQLLLRKAKMRLLYLK---TEDEYDVQSKOMTHLLDQFTNI 353  
604 QRMA-----GMFOVTFEASRLEDEINRTTAEDLPFATSYIVFLVISL 649  
354 KNILALKIEVPGVGGLQGEQKVVHTLSRQLEFEATSVTVIPVFLAYILIPAVT 413  
650 ALGYSYSSWSRVMDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSVLIVQVPLVLSV 709  
414 SCFRPD-----CIRNKMCAAFGVISAFVAVGFLGLHIGV--PFVIVANSFPLILGV 467  
710 GADNIFIVLEYSQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLIGALTPMPAV 769

Db 468 GYDDMFIMISAWHK-TNLAGDIRE-RMSNVYSKAAVSITITITITNIALYTGIMSFPSV 525  
QY 770 RTFALTSGLAVIDFELLQMSAFVALLSLDSKROEASRLDVCCKVPQELP----- 819  
Db 526 QCFCIYTGTTLLFCYFYNYITCFGAFALDQKEV-----VCLWLKADPKWPSFKKFC 580  
QY 820 -----PPQGEGL--LLGFFOKAYAPFLHWHITRGVLLLLFLALPGVSLYSMCHISVGL 871  
Db 581 PFGSVPDDEGTDIHPMSLFFRDYFGPFLTRSESKYFVVFVLYVLISSYIGCFHVQEG 640  
QY 872 DQELALPKDSYLLDYFLNRYF-EVGAPYVFTTILGYNPFSEAGNAICSSAGCNFSE 930  
Db 641 DLNRLASDDSYITPYENVEENYFSDYGRVMVITVKVDYWDK-----DV 685  
QY 931 TQKIQVATEFPQSILAIP--ASSWVDDFIDMLTPSSCCRLYISGNPKDKFCFSTVNSLN 988  
Db 686 RQLENCITKIFKKNVYVDKNTLFEWLDAYVQ-----YLGKNSQD---PNEKNT-- 730  
QY 989 CLKNCMSITMGSVRPSVEQFHKYLPWFLNDRPNKCPKGLAAAYSTSVNLTSGOVLASR 1048  
Db 731 -----FMNIPDLSNFPN-----FOHDINISSNEIISR 761  
QY 1049 FMAYHKPLKNSQDYTEALRAARELANITADIAKVPCTDPAFVFPVPTITNVFVEQVLT 1108  
Db 762 GFQTDTDVSS-----AKKI---LLFQLRRI-AEDCQIPLMVNQAFIYFDQYAAI 809  
QY 1109 LPE-----GLEMLSLCLVP-----TFAVSCLLLGLDLRLSGLNLLSIVMLV 1150  
Db 810 LEDTVNVLVASAMFIVSLLIPYPLCSLWTFVIGSVIVG----- 852  
QY 1151 DTGCFMALWDISYNAVSLINLVSAVGMVSFVSHITSFALSTKPTWLERAKEATISMG 1210  
Db 853 --TCGMAFKVNLDSISMINLVICIGFSDFSAHISAFVSSQSPVQKSVREALYLLGY 910  
QY 1211 AVFAGVANTNLPGLVLTGLAKAQLIQIFFERLNLITLLGLHGLVFLPVLSVYG 1266  
Db 911 PVLOS-AISTTIIGVCVLAARAYIFRT-FFKIMFLVMIFGAHGLIFIPVLTFFG 964

RESULT 13  
US-11-332-764-2  
Sequence 2, Application US/11332764  
GENERAL INFORMATION:  
APPLICANT: Wisotzkey, Robert G.  
TITLE OF INVENTION: PATCHED HOMOLOG 2 (PTCH-2) DISRUPTIONS, COMPOSITIONS AND METHOD  
TITLE OF INVENTION: RELATING THERETO  
FILE REFERENCE: R1673 CIP/75658.064000  
CURRENT APPLICATION NUMBER: US/11/332,764  
CURRENT FILING DATE: 2006-01-13  
PRIOR APPLICATION NUMBER: US 60/413,543  
PRIOR FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: US 10/669,143  
PRIOR FILING DATE: 2003-09-23  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 2  
LENGTH: 1182  
TYPE: PRT  
ORGANISM: Mus musculus  
US-11-332-764-2

Query Match 8.4%; Score 583.5; DB 6; Length 1182;  
Best Local Similarity 24.2%; Pred. No. 4.7e-38;  
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;  
QY 330 SFSTHTLLQO-----PFGQ-----WGTWASWPLTILVSVIPWVALAAGLVFTEL 376  
Db 20 SSAPHTLAGSLQAPLWLRAYFOGLLSLGRICQKCGKVLFLGLVAFGALGLRVAVIE 79  
QY 377 TDPVELWSAPNSQARSEKAFHGHG-PFFRTNQVILTAPNRSRYSDSLLLGPKPFSGI 435  
Db 80 TDLEQLWVEGSRVSQELHYTKEKLBEEAAYTSQMLI-----QTAHOBGGNVLTPE-----A 131



302	TARDLQQLRAEALQSTFLMSPRQLYEHPRGDGYQTHDIGNWSEQASMWLQAWRRFPVQ	361
610	MFQVTFTAERSLE-----DEINRTAEDLPFATSYIV--IFLYISIALGSSYSSWS	658
362	LAQALPANASQQIHAFSSTTTDDILRAFSE-----VSTTRVVGYYLLMLAYACVMTLWRD	417
659	RVWVDSXATLGLGVAVVLGNVMAAGFFSVGLTRSSLVILQVVPFLVLSVGADNFI FV	718
418	--CAQSGAVGLAGVLLVALAVASGLGCLALLGTTFNAATTQVLPFLALGIGVDDIFLLA	475
719	LEYORLPRRQCEPREVHIGRALGVAPSMLLCSLEAICFFLGTALTPMPAVRFTFALTSGL	778
476	HAFTKAP--PDTPLPERMGECIRSTGTISVALTSVNNVAFPMALVPIPALRAEFSLOAAI	533
779	AVILDFFLQWSAFVALLSLOSROEASRLNVCC-----VKPQE-----	817
534	VWGCNFAAVMLVFPAILSLDLRRHRORQLDVLCCFSPCSAQVQIOMQLPQELGDRAVPVGI	593
818	-----LPP-----PG-----QGBG--	826
594	AHLTATVQAFTHCBASSQHVVTILPQAHLLSPASDPLGSELVSPGSGTRDILLSQEBGTG	653
827	-----LLLGFFOKAYAPFLLHMITRGVLLFLFALFGVLSYMSCHISVGLDQ	873
654	PQAAACRPLLCAHWTLAHFARYQAPAPLLQITRAKALVLLFFCALLGLSLYGATLVQDGLAL	713
874	ELALPKDSYLLDYFLNRRYPEVCAPYVFTTILGYNFS--SEAGMNAICSSAGCWNFSFTQ	932
714	TDVVPRGTGKHAFLSAQLRYFSL--YEVALVTQGGFDYAHSQRAL-----FDLHQ	761
933	KIQVATEFFEQSYLAIPAS-----SWVDDF-----IDWLTPSSCCRLYISGPNK	976
762	RFSSL-----KAVLPPATQAPRTHLVHYSWLOGIOAAFDQDWSGREITCHSVRNGSED	816
977	D-----KFCPSTVNSLNCIKNCMSITM-----GSVRPSVEQFHXYLPWFNLDRPNKICPK	1026
817	GALAYKLLIQTGAQEPDLDFSQLTTRTKLVNDKEGLIPP--ELFVGLYVTVVSSDPL-----	869
1027	GGIAA-----YST--SVNLTSQGVLASRFEMAYHKPLKNSODYTEALR	1067
870	GLAASQANFYPPPEWHKDYDTTGENLRIPAAQPLEPAQFPPLHLGLQKTADFVEAIE	928
1068	AARELAANI-TADLRKVPGTDPAPFEPVPTYITNVFYEQYLTILPEGLFMLSCL--LVPTTF	1124
929	GARAACTEAGQAGVHAYPSGSPF-----LFWEQYLGCL--RRCFLLAVCILLVCTFE	976
1125	AVSCILILGLDLRSGLINLLSIVMTLVDTVGFMAWDISYNAVSLINILVSAVGSVEFVSH	1184
977	LVCALLLSLSPWTAGLI-VLWAMTVBELFGIMFLGKLSAIPVWILVASIGIGVEFTVH	1035
1185	ITRFSJAISTKPTWLERAKEATISMGSAVFAGV---AMTNLPGLIIVLGLAKAQITQIFPFR	1241
1036	VALGFLTSHGSRNLRAA-----SALEQTFAPVTDGAVSTLLGLLMLAGSNFDFIIRFFV	1090
1242	LNLLITLGLLHGLVFLPVLISVYGP	1267
1091	VLTVTLGLLHGLLFLPVLISIGP	1116

RESULT 15

US-60-772-265-1197

03-00-772-203-1137  
; Sequence 1197, Application US/60772265

; SEQUENCE 1137, APPRI  
; GENERAL INFORMATION:

APPLICANT: BOUKHAROV, ANDREY A.

APPLICANT: DU, ZIJIN

; APPLICANT: GUO, LIANG

APPLICANT: HRESKO, MICHELLE C

APPLICANT: KOVALIC, DAVI

APPLICANT: ZHAOLONG, LI

; APPLICANT: LU, MAOLONG

APPLICANT: MCCARTER, JAMES P  
APPLICANT: MITCHELL, NANCY M

APPLICANT: MILLER, NANCY  
APPLICANT: VAIDYN, MARK

; APPLICANT: VAUDIN, MARK

```

; APPLICANT: WILLIAMS, DERYCK J
; APPLICANT: WU, WEI
; TITLE OF INVENTION: IDENTIFICATION AND USE OF TARGET GENES
; TITLE OF INVENTION: FOR CONTROL OF PLANT PARASITIC NEMATODES
; FILE REFERENCE: MNDI:002USP1
; CURRENT APPLICATION NUMBER: US/60/772,265
; CURRENT FILING DATE: 2006-02-10
; NUMBER OF SEQ ID NOS: 1919
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1197
; LENGTH: 891
; TYPE: PR2
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: genomic DNA sequence=SeqID_396; coding sequence=SeqID_991
US-60-772-265-1197

```

[illegible]

Qy	1212	VFAGV----	AMTNLPGLVLGLAKAQLIQIEFFRLNLLITLLGLHGLVFLPVILSVGP-	1267
Db	735	MFIPVIHGGLSTLLGIVMLAFSEBDFVVKYFFVVM TALVIIGLGLALLPVLLSLIGPP	794	
Qy	1268	-----	DVNPALALEQKRAEEA VAVVASCNHP-----	SRVSTADNIYV 1307
Db	795	CEITPVNGSNLLPCPASERYRPEDSPNCFTTPSTHQQSD	EQKRRRGCKRMLAVDD---	851
Qy	1308	NHSFEGSIKGAGA	1320	
Db	852	SSSASSTEDGAGA	864	

Search completed: April 7, 2006, 19:33:58  
Job time : 83 secs

**THIS PAGE IS BLANK**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE IS BLANK**